

**Evolutionary Conservation and Characterization of  
the Metazoan Amino Acid Response**

By

Maja Lena Edenius

B.A., Boston University, Biology (2008)

Submitted in partial fulfillment of the requirements for the degree of  
Doctor of Philosophy  
at the  
MASSACHUSETTS INSTITUTE OF TECHNOLOGY  
and the  
WOODS HOLE OCEANOGRAPHIC INSTITUTION  
February 2018

© 2018 Maja Lena Edenius. All rights reserved.

The author hereby grants to MIT and WHOI permission to reproduce and to distribute publicly paper and  
electronic copies of this thesis document in whole or in part in any medium now known or hereafter  
created.

Signature of Author .....  
Joint Program in Oceanography/Applied Ocean Science and Engineering  
Massachusetts Institute of Technology and Woods Hole Oceanographic Institution  
February 5, 2018

Certified by .....  
Dr. Malcolm Whitman  
Professor of Developmental Biology  
Harvard School of Dental Medicine  
Thesis Supervisor

Certified by .....  
Dr. Mark Hahn  
Senior Scientist in Biology  
Woods Hole Oceanographic Institution  
Thesis Supervisor

Accepted by .....  
Dr. Ann M. Tarrant  
Chair, Joint Committee for Biological Oceanography  
Woods Hole Oceanographic Institution



# **Evolutionary Conservation and Characterization of the Metazoan Amino Acid Response**

by  
Maja Lena Edenius

Submitted to the MIT/WHOI Joint Program in Oceanography/Applied Ocean Science and Engineering on  
February 5, 2018 in partial fulfillment of the requirements for the degree of Doctor of Philosophy in  
Marine Biology

## **ABSTRACT**

Signaling pathways that respond to stress and sense nutrient availability are highly conserved throughout eukaryotes. In mammalian cells, these pathways have evolved to regulate immune responses, representing important therapeutic targets. Interestingly, components of these pathways can be found in plants, yeast and nematodes, where they also participate in response to abiotic and biotic stress. The Amino Acid Response (AAR) pathway, an ancient response to the cellular accumulation of uncharged tRNA, is part of the larger Integrated Stress Response (ISR) in mammals. The ISR consists of multiple branches, each one triggered by distinct stresses that produce phospho-eIF2 $\alpha$  signal generation. Each ISR initiating stress results in a unique cellular response due to activation of both the ISR and additional parallel pathway(s) by the initiating stress, but, to date, no such alternate pathway has been identified for the AAR pathway. Despite its integral role in stress adaptation, the ISR has not been studied in early diverging animals. I have identified a highly conserved phosphorylation site in the protein eIF2 $\alpha$ , the signature ISR effector, which allowed me to use a mammalian antibody to identify and characterize the ISR in the basal metazoan, *Nematostella vectensis*, revealing that the core components of the mammalian ISR were present over 550 million years ago in the common ancestor of cnidarians and bilaterians. Additionally, our lab has discovered a novel branch of the AAR pathway that regulates key tissue protective signals. Using evolutionary conservation of this pathway in model organisms, I have identified GCN1 as the branch point that links the signal generation components of the AAR pathway to downstream therapeutic effects. I then used transcriptomic and protein interaction analyses to begin to understand the scope of this pathway and identify key pathway regulators.

Thesis Supervisors:

Malcolm Whitman  
Professor of Developmental Biology  
Harvard School of Dental Medicine

Mark Hahn  
Senior Scientist in Biology  
Woods Hole Oceanographic Institution

Thesis Advisors:

Professor Malcolm Whitman  
Department of Developmental Biology  
Harvard School of Dental Medicine

Dr. Mark Hahn  
Department of Biology  
Woods Hole Oceanographic Institution

Thesis Committee:

Professor Dennis Kim  
Department of Biology  
Massachusetts Institute of Technology

Professor Gerald Thomsen  
Department of Biochemistry and Cell Biology  
Stony Brook University

Thesis Defense Chair:

Professor Martin Polz  
Department of Civil and Environmental Engineering  
Massachusetts Institute of Technology

## ACKNOWLEDGEMENTS

This thesis wouldn't have been possible without my incredible and supportive mentors, colleagues, friends, and family. It would be impossible to thank all the people who have helped and guided me through the last five and a half years and below is just an attempt to begin to thank some of these people, who have together made this adventure worthwhile and possible.

Thank you to my advisor, Malcolm Whitman for showing me how to be a scientist. For teaching me to see the important details in each experiment and to not just follow the accepted methods and ideas, but to consider them for myself and explore all the possibilities. Thank you for always having your door open, no matter how busy you were, and no matter how many questions I had. For always being interested and available to discuss my experiments, whether it was how to best extract protein from sea anemones or how to isolate ribosomes from human synoviocytes. Thank you for teaching me how to be a good scientist, both as a biologist and a person, and for giving my sea anemones a home in a dental school.

To my mentor, Tracy Keller, for teaching me how to think, write, and talk about science. You make science fun and exciting! Thank you for all the interesting discussions over the years, and above all for always believing in me. For supporting my interests in marine biology and discussing and developing ideas on anything ranging from sea stars to pharmaceuticals. You are a true scientist, able to ask and develop ideas even in areas not your area of expertise. I hope a little bit has rubbed off on me.

To my co-advisor, Mark Hahn. Thank you for all the data discussions and for keeping me on track. No matter how confusing my data was you always cut straight to the important facts and details, providing important insight and advice. Thank you for the patient support and encouragement. To my thesis committee, Dennis Kim and Gerald Thomsen, for your input and support, and for all the traveling and phone calls!

And I couldn't have done this without my "magical" lab. To Chang-Yeol Yeo for teaching me cell biology. Thank you for all your help over the years, you were introduced to me as "the cell biology god" and you have lived up to your name. To Yeon-Jin Kim whose work has been integral to the work of my thesis. Thank you for your help, always with a smile, and a helpful hand. To Kristen Powers and Leila Revollo for the many motivational sticky notes and unicorns. Couldn't have done it without the sticky notes! An additional thanks to Kristen, for helping me with ordering and caring for my cells when I was at WHOI. To Minjin Bahn for the much needed coffee trips and Changqian Zhao for finally getting our CRISPRs to work. And finally, to Mattia Bordoli for keeping the lab on schedule and answering my many, many questions over the years.

To my classmates and WHOI colleagues, in particular, Amy Maas, Amalia Aruda Almada, and Harriet Alexander, for your scientific and moral support.

To my friends, who will remain nameless here, thank you for the encouragement and distractions. I am very lucky to have you all. Thank you for keeping me happy!

And finally, tack ja kiitos to my family, for teaching me that nature and the natural world are not separate from us, but an integral part of our lives. And thank you aiti for the Finnish sisu.

-----

Funding for this research was provided by the National Science Foundation Graduate Research Fellowship Program, Allied Bristol Life Sciences (to Malcolm Whitman), the WHOI Academic Programs Office, and the WHOI Ocean Venture Fund.



## TABLE OF CONTENTS

<b>Abstract.....</b>	<b>3</b>
<b>Acknowledgements .....</b>	<b>5</b>
<b>Chapter One: Introduction.....</b>	<b>9</b>
<b>Figures.....</b>	<b>15</b>
<b>References.....</b>	<b>17</b>
<b>Chapter Two: Characterization of Integrated Stress Response (ISR) Pathway and its Activation in the Basal Metazoan Lineage, <i>Nematostella vectensis</i>.....</b>	<b>21</b>
<b>Tables and Figures.....</b>	<b>35</b>
<b>References.....</b>	<b>46</b>
<b>Chapter Three: The Evolutionarily Conserved Protein GCN1 is a Novel Branch Point in the Mammalian Amino Acid Response (AAR) Pathway.....</b>	<b>53</b>
<b>Tables and Figures.....</b>	<b>67</b>
<b>References.....</b>	<b>77</b>
<b>Chapter Four: Characterization of a Novel Mammalian Nutritional Stress Response Pathway.....</b>	<b>81</b>
<b>Tables and Figures.....</b>	<b>89</b>
<b>References.....</b>	<b>96</b>
<b>Chapter Five: Conclusions .....</b>	<b>97</b>
<b>References.....</b>	<b>101</b>





## **CHAPTER ONE**

### **Introduction**



## I. All cells respond to nutrients and stress

All cells have the capacity to sense and respond to factors in their environment, particularly nutrients, required to generate the energy and building blocks of the cell. Cells and organisms must coordinate their growth with nutrient availability. Many nutrient sensing pathways of the cell are also co-opted for cellular stress response, mediating adaptive and protective changes not only to nutrients but to other abiotic and biotic factors (1-7). In mammals, these pathways have further evolved functions in immunity (8-12). Many constituents of nutrient and cellular stress response pathways are among the most highly evolutionarily conserved genes across all three domains of life (13). Although much of this ancient cellular machinery has been studied extensively, some basic components of highly conserved nutrient response pathways remain poorly understood.

In eukaryotic cells, two highly conserved pathways are known that sense and respond to intracellular amino acid levels, mechanistic Target Of Rapamycin (mTOR)C1 and the Amino Acid Response (AAR) pathway. Both pathways regulate adaptive changes to maintain homeostasis in response to amino acid supply. Interestingly these pathways are key regulators of lifespan in model organisms and both pathways regulate immune function in mammalian cells (14-17). Although both pathways respond to amino acid levels, they are differentially activated and regulate distinct cellular programs. Unlike the AAR pathway, the mTORC1 pathway is thought to sense amino acid levels directly and is inhibited by amino acid restriction (18). The AAR pathway, in contrast, is activated by intracellular accumulation of uncharged tRNAs that can result from the lack of any amino acid or from inhibition of any aminoacyl-tRNA synthetase (19). Uncharged tRNA binds and activates the kinase GCN2, resulting in phosphorylation of eIF2 $\alpha$ , a shared component of multiple stress response pathways, referred to as the Integrated Stress Response (ISR). Study of the AAR pathway, conserved across eukaryotes, led to the discovery of the expanded mammalian integrated stress response (ISR), a group of stress sensing pathways that converge on phosphorylation and activation of eIF2 $\alpha$ , a master regulator of stress response in mammalian and likely all metazoan cells (20, 21).

In this thesis, one goal was to use the high degree of conservation of components of the ISR and AAR pathway to determine whether the core mammalian ISR is conserved in the basal metazoan cnidarian, *Nematostella vectensis*. The ISR's broad conservation within eukaryotes suggests that it likely plays a similar role in animals from early diverging lineages such as cnidarians (1, 22). Although it is clear that cnidarians respond to environmental stressors, little is known about the molecular signaling pathways that evoke these responses (23, 24). Signaling pathways regulating cellular stress responses have been extensively studied in classic model eukaryotes such as yeast, mice, and humans, but few studies have explored these pathways in early diverging animal lineages. In addition, I used the evolutionary conservation of the AAR pathway throughout eukaryotes to elucidate an unknown signaling pathway triggered by uncharged tRNA observed by our lab in mammalian cells. Our lab previously solved the mechanism of the small molecule therapeutic halofuginone (HF) (**Fig 1**) by demonstrating that it binds to and inhibits the prolyl-tRNA synthetase (PRS) activity of the fused amino-acyl tRNA synthetase (aaRS) glutamyl-prolyl tRNA synthetase (EPRS) (25, 26). Because HF is a potent inhibitor of PRS, it acts as an amino acid restriction mimetic, and, like other aaRS- inhibitors, it triggers the cellular accumulation of uncharged tRNAs thereby activating the AAR pathway (25). Our lab found that in an *in vitro* system of TNF $\alpha$ -stimulated fibroblast-like-synoviocytes (FLS), cells implicated in the pathology of rheumatoid arthritis, treatment with either HF or the threonyl-tRNA synthetase inhibitor, borrelidin, suppresses a pro-inflammatory transcriptional program in these cells (Kim et al, manuscript in preparation). Likewise, our scientific collaborator Mark Sundrud has found that HF or borrelidin treatment inhibits *in vitro* differentiation and proliferation of pro-inflammatory Th17 cells (27). Our lab and the Sundrud lab further made the key observation that these downstream inflammatory suppressive effects triggered by uncharged tRNA occurs in cells that lack the signature AAR pathway effector, GCN2. GCN2 is considered the sole sensor of uncharged tRNA in the eukaryotic cell, but these data suggest there is an undescribed

mammalian signaling pathway that responds to amino acid limitation, and an unknown sensor of uncharged tRNA.

## **II. All cells have evolved mechanisms to sense stress through the translational apparatus**

### **Translation is an expensive process regulated in all cells**

All cells regulate protein translation. From microbes to higher organisms the availability of amino acids either endogenously synthesized or acquired from the environment is necessary for survival (28). Amino acids are the building blocks of proteins, which are essential for the structure, function, and regulation of the cell; however, protein synthesis is an energetically expensive process, using 22-30% of a cell's total energy (3, 29). Failure to finish peptide chain synthesis due to stalled ribosomes under amino acid scarcity is energetically inefficient and cells primarily regulate translation at the initiation step (29, 30). To keep translation in tune with cellular conditions, all cells have evolved mechanisms to sense stress through the translational apparatus.

## **III. The Bacterial Stringent Response (Fig 2)**

The bacterial stringent response is a paradigm of stress response, which results in massive transcriptional reprogramming (31). Upon amino acid deprivation, stalled ribosomes containing uncharged tRNA at the A-site accumulate. The enzyme, RelA detects and binds uncharged tRNA and stalled ribosomes at the A-site, where RelA then synthesizes a hyper phosphorylated guanine analogue, (p)ppGpp, from GTP/GDP and ATP (31). ppGpp acts as a master regulator of gene transcription, upregulating adaptive genes like amino acid biosynthetic genes to counteract low amino acid levels and protective or defensive genes like virulence factors. ppGpp directly regulates transcription by binding the RNA polymerase core enzyme. In addition, ppGpp binds bacterial initiation factor 2, decreasing its binding to fMet-tRNA<sup>fMet</sup>, downregulating protein translation to conserve energy (31-34). ppGpp not only increases during amino acid starvation, but also in response to other types of nutrient limitation and factors causing growth arrest (32). The stringent response plays an important role in antibiotic resistance and survival of pathogenic bacteria under stress. The stringent response has been found in bacteria and plant plastids, but not in archaea, yeast, or animals (32, 35, 36). ppGpp has not been found in archaea, but an eIF2 $\alpha$  kinase homolog has been identified (37, 38).

## **IV. The Eukaryotic Amino Acid Response (AAR) Pathway**

Although the stringent response has not been found in eukaryotes, other than plant plastids, a eukaryotic pathway that shares the hallmarks of the stringent response is the AAR pathway, in which uncharged tRNA signals amino acid depletion at the ribosome resulting in translational inhibition and adaptive and protective gene expression. The AAR pathway was first discovered in yeast over 40 years ago, in which it was observed that a large number of genes encoding enzymes in multiple amino acid biosynthetic pathways were under common control (39, 40). In *Saccharomyces cerevisiae*, starvation for an amino acid or the presence of a defective aminoacyl-tRNA synthetase leads to increased transcription of more than 30 genes encoding amino acid biosynthetic genes. Since this response leads to enhanced transcription of amino acid biosynthetic genes from several different amino acid biosynthetic pathways, the response was called the general amino acid control (GAAC) (40). Genetic analyses identified several *trans*-acting positive factors required for the GAAC including general control nonderepressible or GCN1, GCN2, GCN3, and GCN4 that constitute the major components of the AAR pathway (40).

### **The AAR Pathway: the GCN2-eIF2 $\alpha$ -ATF4 pathway (Fig 3a)**

Accumulation of uncharged tRNAs, resulting from amino acid restriction or aaRS-inhibition, is sensed at the ribosome. The kinase GCN2, bound to the ribosome, binds uncharged tRNA, resulting in autophosphorylation and phosphorylation of the alpha subunit of the eukaryotic initiation factor 2 (eIF2 $\alpha$ )

(GCN3 of the yeast mutant analyses) (41). GCN2 is considered a tRNA sensor since it binds tRNA through its histidyl-tRNA synthetase domain, and has been shown to bind uncharged tRNA with a 5-fold preference over charged tRNA *in vitro* (41, 42). eIF2 is central to translation and mediates binding of initiator methionyl-tRNA to ribosomes, but phosphorylation of its alpha subunit inhibits recycling of the eIF2-MetTRNA-GTP ternary complex inhibiting cap-dependent protein translation (43). Paradoxically, phosphorylation of eIF2 $\alpha$  results in increased translation of specific mRNAs that contain upstream open reading frames in their 5' leader including the bZIP transcription factor, GCN4 in yeast, or its functional equivalent in vertebrates, ATF4 (activating transcription factor 4), which sets in motion cascades of gene expression involved in amino acid metabolism and transport, redox homeostasis, and signaling (44-48).

### **The AAR pathway is one of four ISR arms in mammalian cells (Fig 3b)**

In mammals the GCN2-eIF2 $\alpha$ -ATF4 pathway has evolved to respond to a larger range of stressors with additional eIF2 $\alpha$  kinases that each respond to distinct signals, and is referred to as the integrated stress response (ISR) (49). The ISR has primarily been studied in mammals and the extent to which the ISR is conserved across metazoa remains unclear. Induction of the ISR by different stressors results in shared effects because all eIF2 $\alpha$  kinases converge on eIF2 $\alpha$ , but each activator kinase arm also has independent transcriptional readouts and functions (50-52). The AAR pathway is the sole ISR arm thought to signal through eIF2 $\alpha$  alone. Our data now suggest that the accumulation of uncharged tRNA, like other ISR stress inputs, has signal readouts other than eIF2 $\alpha$ .

### **HF links nutritional stress response to tissue protection**

Our lab recently discovered the cellular target of halofuginone (HF), a semi-synthetic derivative of the natural product febrifugine, by showing that it binds EPRS, inhibiting its PRS activity (25). PRS is the cellular enzyme responsible for charging proline into its cognate tRNA. Inhibitors of PRS or any aaRS act as amino acid restriction mimetics, activating the mammalian AAR pathway (19, 53). The plant alkaloid febrifugine, found in the roots of blue evergreen hydrangea, *Dichroa febrifuga* Lour, is the active ingredient of an ancient Chinese herbal remedy that has been used for roughly 2000 years for its therapeutic effects against several ailments including malaria and fever (54). Its traditional use for malaria led to investigations to evaluate if it could be used to develop a new antimalarial drug. The attempt failed due to potent side effects, but a search for a less toxic form in the late 1960s led to the design of the racemic halogenated derivative HF (55). Even prior to an understanding of its tissue mechanism, HF progressed to phase II clinical trials for fibrotic disease, bladder cancer, and Duchenne muscular dystrophy (25, 56-59). Discovery of the HF target now links a variety of therapeutic effects, including anti-fibrotic, anti-inflammatory, anti-angiogenic, and anti-metastatic activities to tissue induction of nutrient stress response pathways (26, 58-61).

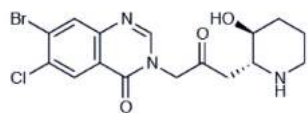
## **V. A novel nutritional stress response pathway**

The AAR pathway in eukaryotic cells shares hallmarks of the stringent response, in which uncharged tRNA signals amino acid depletion at the ribosome resulting in translational inhibition and adaptive and protective gene expression. Our lab's data point to the existence of an additional pathway in mammalian cells, in which the production of uncharged tRNA triggers a broad transcriptional cellular reprogramming to dampen inflammation and protect host tissues. GCN2, the signature mediator of the AAR pathway, is the only protein known to sense and signal in response to uncharged tRNA in mammalian cells. Studies from our lab using the aaRS-inhibitors HF and borrelidin to induce a tissue protective program in FLS; however, clearly show that this pathway is intact in cells lacking GCN2 (Kim et al, manuscript in preparation). The evolutionary conservation of the AAR pathway provides hints for where to start looking to uncover this unidentified pathway. To identify candidate mediators of the novel tissue protective pathway, I looked to the evolutionary conservation of the AAR pathway. GCN1 is a protein conserved throughout eukaryotes, whose function has only been studied in yeast. In yeast, GCN1 forms a complex with GCN2, and is genetically required for GCN2 activation by uncharged tRNA (Fig 3c) (62-66). A

large scaffold-like protein with no known enzymatic function, it is plausible GCN1 co-detects uncharged tRNA alongside GCN2, mediating the uncharged tRNA signal not only to GCN2 but to GCN2-independent pathways. In Chapter 3, I will study GCN1 as candidate mediator of the HF-driven GCN2-independent effects observed by our lab, and test the possibility that GCN1 is a branch point in the mammalian AAR pathway. GCN1 provides a candidate mediator of the GCN2 independent effects of uncharged tRNA observed by our lab, representing a possible branch point within the AAR, which will be elucidated in my thesis. For my doctoral thesis I will address the following questions:

- 1. Is the mammalian ISR conserved in the basal metazoan, *Nematostella vectensis*?**
- 2. Is GCN1 a branch point in the mammalian AAR pathway?**
- 3. What is the scope of the transcriptional program regulated by the novel HF-activated pathway, and what effectors mediate these GCN1-dependent GCN2-independent signals?**

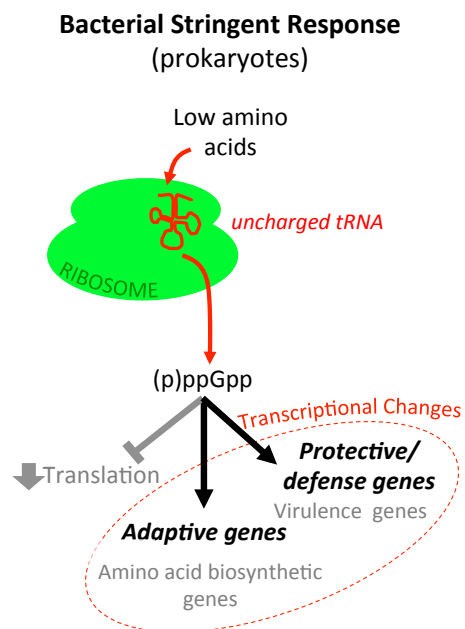
## FIGURES



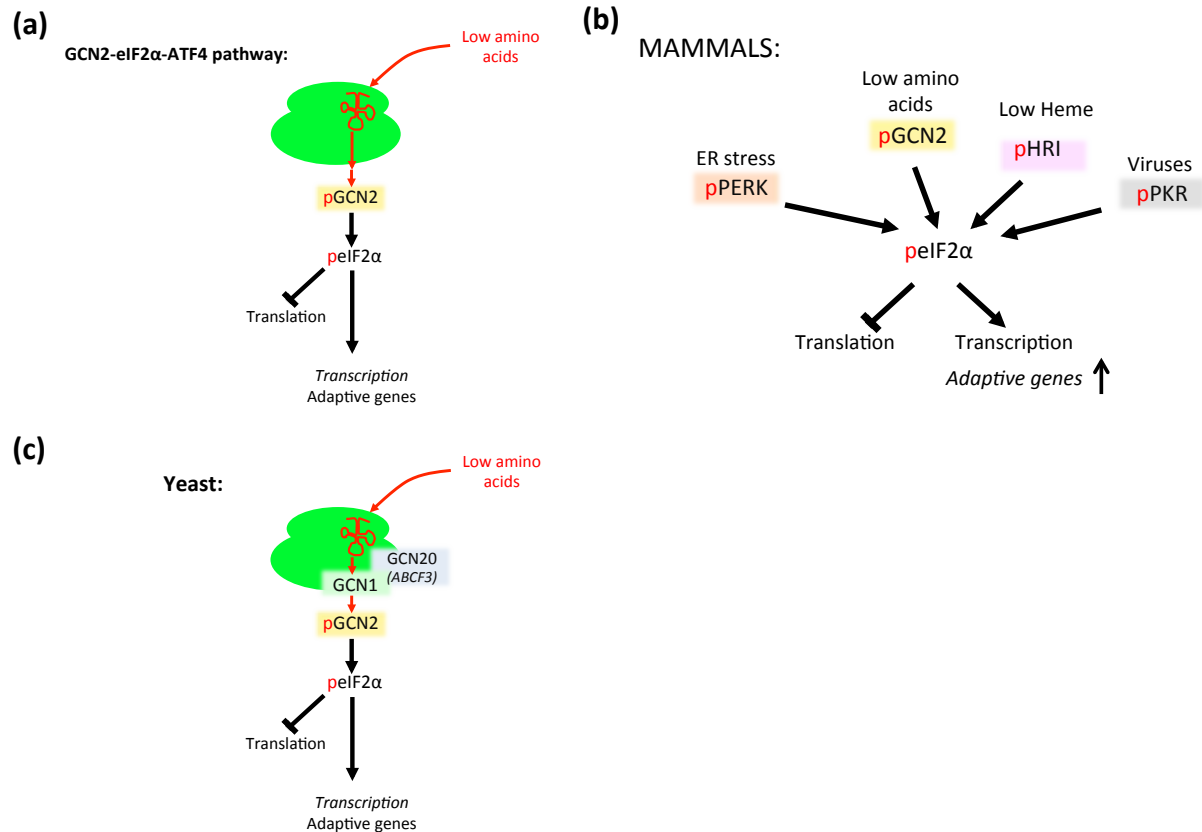
**Halofuginone**

7-bromo-6-chloro-3-[(3-hydroxy-2-piperidinyl)-2-oxopropyl]-4(3H)-quinazolinone

**Fig 1 | Structure of Halofuginone**



**Fig 2 | Cartoon of the bacterial stringent response.**



**Fig 3 | The Amino Acid Response (AAR) pathway and the Integrated Stress Response (ISR).** (a) Cartoon of the AAR pathway. (b) Cartoon of the ISR. (c) Cartoon of the GCN2 pathway in yeast. In yeast the GCN1-GCN20 complex functions upstream of GCN2 and is required for GCN2 activation by uncharged tRNA.



## REFERENCES

1. Lageix S, *et al.* (2008) Arabidopsis eIF2alpha kinase GCN2 is essential for growth in stress conditions and is activated by wounding. *BMC plant biology* 8:134.
2. Leung GP, Aristizabal MJ, Krogan NJ, & Kobar MS (2014) Conditional Genetic Interactions of RTT107, SLX4, and HRQ1 Reveal Dynamic Networks upon DNA Damage in *S. cerevisiae*. *G3: Genes | Genomes | Genetics* 4:1059-1069.
3. Wang L, *et al.* (2017) The inhibition of protein translation mediated by AtGCN1 is essential for cold tolerance in Arabidopsis thaliana. *Plant, cell & environment* 40(1):56-68.
4. Zarattini M & Forlani G (2017) Toward Unveiling the Mechanisms for Transcriptional Regulation of Proline Biosynthesis in the Plant Cell Response to Biotic and Abiotic Stress Conditions. *Frontiers in plant science* 8:927.
5. Zeng W, *et al.* (2011) A genetic screen reveals Arabidopsis stomatal and/or apoplastic defenses against Pseudomonas syringae pv. tomato DC3000. *PLoS pathogens* 7(10):e1002291.
6. Luna E, *et al.* (2014) plant perception of b-aminobutyric acid is mediated by an aspartyl-trna synthetase. *Nature Chemical Biology* 10:450-456.
7. Schwarzenbacher RE, Luna E, & Ton J (2014) The discovery of the BABA receptor: scientific implications and application potential. *Frontiers in plant science* 5:1-3.
8. Walls J, Sinclair L, & Finlay D (2016) Nutrient sensing, signal transduction and immune responses. *Seminars in immunology* 28(5):396-407.
9. Grohmann U, *et al.* (2017) Amino-acid sensing and degrading pathways in immune regulation. *Cytokine & growth factor reviews* 35:37-45.
10. Munn DH & Mellor AL (2013) Indoleamine 2,3 dioxygenase and metabolic control of immune responses. *Trends in immunology* 34(3):137-143.
11. Cobbold SP, *et al.* (2009) Infectious tolerance via the consumption of essential amino acids and mTOR signaling. *Proceedings of the National Academy of Sciences of the United States of America* 106(29):12055-12060.
12. Thomson AW, Turnquist HR, & Raimondi G (2009) Immunoregulatory functions of mTOR inhibition. *Nature reviews. Immunology* 9(5):324-337.
13. Kultz D (2003) Evolution of the cellular stress proteome: from monophyletic origin to ubiquitous function. *Journal of Experimental Biology* 206(18):3119-3124.
14. Johnson SC, Rabinovitch PS, & Kaeberlein M (2013) mTOR is a key modulator of ageing and age-related disease. *Nature* 493(7432):338-345.
15. Ferraz RC, *et al.* (2016) IMPACT is a GCN2 inhibitor that limits lifespan in Caenorhabditis elegans. *BMC biology* 14(1):87.
16. Fontana L, Partridge L, & Longo VD (2010) Extending Healthy Life Span-From Yeast to Humans. *Science* 328:321-326.
17. Xiao F, *et al.* (2011) Leucine Deprivation Increases Hepatic Insulin Sensitivity via GCN2/mTOR/S6K1 and AMPK Pathways. *Diabetes* 60:746-756.
18. Wolfson RL & Sabatini DM (2017) The Dawn of the Age of Amino Acid Sensors for the mTORC1 Pathway. *Cell metabolism* 26(2):301-309.
19. Kilberg MS, Pan YX, Chen H, & Leung-Pineda V (2005) Nutritional control of gene expression: how mammalian cells respond to amino acid limitation. *Annual review of nutrition* 25:59-85.
20. Harding HP, *et al.* (2003) An Integrated Stress Response Regulates Amino Acid Metabolism and Resistance to Oxidative Stress. *Molecular cell* 11:619-633.
21. Taniuchi S, Miyake M, Tsugawa K, Oyadomari M, & Oyadomari S (2016) Integrated stress response of vertebrates is regulated by four eIF2alpha kinases. *Scientific reports* 6:32886.
22. Hernandez G, Altmann M, & Lasko P (2010) Origins and evolution of the mechanisms regulating translation initiation in eukaryotes. *Trends in biochemical sciences* 35(2):63-73.

23. Tarrant AM, Reitzel AM, Kwok CK, & Jenny MJ (2014) Activation of the cnidarian oxidative stress response by ultraviolet light, polycyclic aromatic hydrocarbons and crude oil. *The Journal of experimental biology*.
24. Reitzel AM, *et al.* (2013) Physiological and developmental responses to temperature by the sea anemone *Nematostella vectensis*. *Marine Ecology Progress Series* 484:115-130.
25. Keller TL, *et al.* (2012) Halofuginone and other febrifugine derivatives inhibit prolyl-tRNA synthetase. *Nature Chemical Biology* 8:311-317.
26. Zhou H, Sun L, Yang XL, & Schimmel P (2013) ATP-directed capture of bioactive herbal-based medicine on human tRNA synthetase. *Nature* 494(7435):121-124.
27. Sundrud MS, *et al.* (2009) Halofuginone inhibits Th17 cell differentiation by activating the amino acid starvation response. *Science* 324:1334-1338.
28. Grohmann U & Bronte V (2010) Control of immune response by amino acid metabolism. *Immunological Reviews* 236:243-264.
29. Efeyan A, Comb WC, & Sabatini DM (2015) Nutrient-sensing mechanisms and pathways. *Nature* 517(7534):302-310.
30. Sesma A, Castresana C, & Castellano MM (2017) Regulation of Translation by TOR, eIF4E and eIF2alpha in Plants: Current Knowledge, Challenges and Future Perspectives. *Frontiers in plant science* 8:644.
31. Brown A, Fernandez IS, Gordiyenko Y, & Ramakrishnan V (2016) Ribosome-dependent activation of stringent control. *Nature* 534(7606):277-280.
32. Braeken K, Moris M, Daniels R, Vanderleyden J, & Michiels J (2006) New horizons for (p)ppGpp in bacterial and plant physiology. *Trends in microbiology* 14(1):45-54.
33. Magnusson LU, Farewell A, & Nystrom T (2005) ppGpp: a global regulator in *Escherichia coli*. *Trends in microbiology* 13(5):236-242.
34. Srivatsan A & Wang JD (2008) Control of bacterial transcription, translation and replication by (p)ppGpp. *Current opinion in microbiology* 11(2):100-105.
35. van der Biezen EA, Sun J, Coleman MJ, Bibb MJ, & Jones JD (2000) Arabidopsis RelA/SpoT homologs implicate (p)ppGpp in plant signaling. *Proceedings of the National Academy of Sciences of the United States of America* 97(7):3747-3752.
36. Sugliani M, *et al.* (2016) An Ancient Bacterial Signaling Pathway Regulates Chloroplast Function to Influence Growth and Development in Arabidopsis. *The Plant cell* 28(3):661-679.
37. Cellini A, *et al.* (2004) Stringent control in the archaeal genus *Sulfolobus*. *Research in microbiology* 155(2):98-104.
38. Ray WK, Potters MB, Haile JD, & Kennelly PJ (2015) Activation of SsoPK4, an Archaeal eIF2alpha Kinase Homolog, by Oxidized CoA. *Proteomes* 3(2):89-116.
39. Schurch A, Miozzari J, & Hutter R (1974) Regulation of Tryptophan Biosynthesis in *Saccharomyces cerevisiae*: Mode of Action of 5-Methyl-Tryptophan and 5-Methyl-Tryptophan-Sensitive Mutants. *Journal of Bacteriology* 117(3):1131-1140.
40. Hinnebusch AG (2005) Translational Regulation of GCN4 and the General Amino Acid Control of Yeast. *Annu Rev Microbiol* 59:407-450.
41. Dong J, Qiu H, Garcia-Barrio M, Anderson J, & Hinnebusch AG (2000) Uncharged tRNA Activates GCN2 by Displacing the Protein Kinase Moiety from a Bipartite tRNA-Binding Domain. *Molecular cell* 6:269-279.
42. Zhu S, Sobolev AY, & Wek R (1996) Histidyl-tRNA Synthetase-related Sequences in GCN2 Protein Kinase Regulate in Vitro Phosphorylation of eIF-2. *The Journal of biological chemistry* 271:24989-24994.
43. Baird TD & Wek RC (2012) Eukaryotic initiation factor 2 phosphorylation and translational control in metabolism. *Advances in nutrition* 3(3):307-321.
44. Ameri K & Harris AL (2008) Activating transcription factor 4. *The international journal of biochemistry & cell biology* 40(1):14-21.

45. Harding HP, *et al.* (2000) Regulated Translation Initiation Controls Stress-Induced Gene Expression in Mammalian Cells. *Molecular cell* 6:1099-1108.
46. Kilberg MS, Shan J, & Su N (2009) ATF4-dependent transcription mediates signaling of amino acid limitation. *Trends in endocrinology and metabolism: TEM* 20(9):436-443.
47. Lu PD, Harding HP, & Ron D (2004) Translation reinitiation at alternative open reading frames regulates gene expression in an integrated stress response. *Journal of Cell Biology* 167(1):27-33.
48. Hinnebusch AG & Natarajan K (2002) Gcn4p, a Master Regulator of Gene Expression, Is Controlled at Multiple Levels by Diverse Signals of Starvation and Stress. *Eukaryotic cell* 1:22-32.
49. Donnelly N, Gorman AM, Gupta S, & Samali A (2013) The eIF2alpha kinases: their structures and functions. *Cellular and molecular life sciences : CMLS* 70(19):3493-3511.
50. Wek RC & Cavener DR (2007) Translational control and the unfolded protein response. *Antioxidants & redox signaling* 9(12):2357-2371.
51. Furuyama K, Kaneko K, & Vargas V. PD (2007) Heme as a Magnificent Molecule with Multiple Missions: Heme Determines Its Own Fate and Governs Cellular Homeostasis. *Journal of Experimental Medicine* 213:1-16.
52. Jacobs BL & Langland JO (1996) When Two Strands Are Better Than One: The Mediators and Modulators of the Cellular Responses to Double-Stranded RNA. *Virology* 219:339-349.
53. Brown MV, Reader JS, & Tzima E (2010) Mammalian aminoacyl-tRNA synthetases: cell signaling functions of the protein translation machinery. *Vascular pharmacology* 52(1-2):21-26.
54. Herman JD, *et al.* (2015) The cytoplasmic prolyl-tRNA synthetase of the malaria parasite is a dual-stage target of febrifugine and its analogs. *Science Translational Medicine* 7(288).
55. McLaughlin NP, Evans P, & Pines M (2014) The chemistry and biology of febrifugine and halofuginone. *Bioorganic & medicinal chemistry* 22(7):1993-2004.
56. Koon HB, *et al.* (2011) Phase II AIDS Malignancy Consortium trial of topical halofuginone in AIDS-related Kaposi sarcoma. *Journal of acquired immune deficiency syndromes* 56(1):64-68.
57. Nagler A & Pines M (1999) Topical treatment of cutaneous chronic graft versus host disease with halofuginone: a novel inhibitor of collagen type I synthesis. *Transplantation* 68:1806-1809.
58. Pines M & Nagler A (1998) Halofuginone: A Novel Antifibrotic Therapy. *General Pharmacology* 30(4):445-450.
59. Elkin M, *et al.* (1999) Inhibition of Matrix Metalloproteinase-2 Expression and Bladder Carcinoma Metastasis by Halofuginone. *Clinical Cancer Research* 5:1982-1988.
60. Gnainsky Y, *et al.* (2007) Gene expression during chemically induced liver fibrosis: effect of halofuginone on TGF-beta signaling. *Cell and tissue research* 328(1):153-166.
61. Park MK, *et al.* (2014) Halofuginone ameliorates autoimmune arthritis in mice by regulating the balance between Th17 and Treg cells and inhibiting osteoclastogenesis. *Arthritis & rheumatology* 66(5):1195-1207.
62. Marton MJ, Crouch D, & Hinnebusch AG (1993) GCN1, a Translational Activator of GCN4 in *Saccharomyces cerevisiae*, is required for Phosphorylation of Eukaryotic Translation Initiation Factor 2 by Protein Kinase GCN2. *Molecular and cellular biology* 13(6):3541-3556.
63. Garcia-Barrio M, Dong J, Ufano S, & Hinnebusch AG (2000) Association of GCN1-GCN20 regulatory complex with the N-terminus of eIF2alpha kinase GCN2 is required for GCN2 activation. *The EMBO journal* 19(8):1887-1899.
64. Sattlegger E & Hinnebusch AG (2000) Separate domains in GCN1 for binding protein kinase GCN2 and ribosomes are required for GCN2 activation in amino acid-starved cells. *The EMBO journal* 19(23):6622-6633.
65. Kubota H, Ota K, Sakaki Y, & Ito T (2001) Budding yeast GCN1 binds the GI domain to activate the eIF2alpha kinase GCN2. *The Journal of biological chemistry* 276(20):17591-17596.
66. Sattlegger E & Hinnebusch AG (2005) Polyribosome binding by GCN1 is required for full activation of eukaryotic translation initiation factor 2 {alpha} kinase GCN2 during amino acid starvation. *The Journal of biological chemistry* 280(16):16514-16521.



## **CHAPTER TWO**

### **Characterization of Integrated Stress Response (ISR) Pathway and its Activation in the Basal Metazoan Lineage, *Nematostella vectensis***



## INTRODUCTION

### **The Integrated Stress Response (ISR) and the phylogenetic position of *Nematostella***

Control of protein translation is a central component of adaptation to environmental stress. Understanding the mechanisms that underlie cellular adaptation to stress is necessary for the ability to monitor environmental health, predict novel hazardous chemicals, and predict how organisms will respond to toxicants and changing environmental conditions. Translational changes in response to stimuli modify the cellular proteome more rapidly than do transcriptional changes, and provide a first line of defense for the cell (1-3). In addition, these translational changes regulate many of the subsequent transcriptional changes that control longer-term adaptations to stress (4-7). In the cell, translation is primarily regulated at the point of translation initiation, and one of the primary proteins regulating translation initiation is the eukaryotic initiation factor 2 (eIF2) (8). eIF2 is the central component of the integrated stress response (ISR), a signaling pathway that receives and integrates inputs from a range of stressors to mediate protective cellular effects via phosphorylation of eIF2.

Cellular adaptation to stress in response to eIF2 phosphorylation is conserved from unicellular eukaryotes to mammals, but there is a wide range of variation in upstream pathway activators, downstream mediators and transcriptional responses between yeast and mammals (9, 10). The ISR's broad conservation within eukaryotes suggests that it likely plays a similar role in animals from early diverging lineages such as cnidarians (11, 12). Although it is clear that cnidarians respond to environmental stressors, little is known about the molecular signaling pathways that evoke these responses (13, 14). Signaling pathways regulating cellular stress responses have been extensively studied in classic model eukaryotes such as yeast, mice, and humans, but few studies have explored these pathways in early diverging animal lineages. *Nematostella vectensis* has become an increasingly popular cnidarian model in which to study questions about metazoan evolution and evolutionary development, in part due to its key phylogenetic position (15). Cnidarians are an early branching metazoan lineage and in phylogenetic terms are the sister group to Bilateria (16). The split between Bilateria and Cnidaria is estimated to have occurred about 600 million years ago (17). *Nematostella* belongs to class Anthozoa, which includes sea anemones and reef-building corals. The genome of *Nematostella* has been sequenced and shows a slow rate of sequence evolution evidenced by a high level of conservation of ancestral genetic traits (18, 19). *Nematostella* therefor represents a good model to capture ancient functions of conserved genes and signaling pathways within metazoa (20). Phylogenetic position, ease of laboratory culture, and its ecologically important and accessible habitat make *Nematostella* an interesting model for understanding the regulation of stress response in a cnidarian.

The ISR is a broadly conserved regulatory event activated in response to diverse stressors, and exploring the consequences of this event has provided insight into the ways in which stress is sensed and responded to in classical eukaryotic models (21). In this chapter, I identify key components of the ISR in *Nematostella*, which constitutes the first report of this functional stress response pathway in a non-bilaterian metazoan. *Nematostella* is now the most evolutionarily distant animal from humans that has been shown to have a conserved ISR. Studying stress-triggered ISR pathway activation in *Nematostella* will help to understand the response and adaptation of this organism and related organisms to relevant marine stressors. Because of its key phylogenetic position, characterization of the ISR components in *Nematostella* could yield valuable insights into the roots and emergence of the more complex mammalian ISR.

### **Evolutionary conservation and divergence of the ISR**

In the cell, translation is primarily regulated at the point of translation initiation. eIF2 is a key regulator of translation initiation and the central component of the ISR (22-25). A ternary complex that consists of

eIF2, GTP, and methionyl-tRNA recognizes methionine start codons. Phosphorylation of the  $\alpha$  subunit of eIF2 (eIF2 $\alpha$ ) inhibits the recycling of the ternary complex, and it is this phosphorylation of eIF2 $\alpha$  (at Ser51) by ISR kinases that makes eIF2 the central mediator of translational regulation by the ISR. This translational control-based signaling cassette is conserved across eukaryotes and plays a central role in cellular adaptation to stress.

The predominant effect of ISR pathway activation is a general reduction in the initiation of translation, which thereby reduces cell energy use and unfolded, or damaged, protein load. Paradoxically, ISR pathway activation also up regulates synthesis of a small subset of proteins that are required for adaptation to stress, but that aren't translated efficiently in an unstressed cell. Proteins that contain open reading frames upstream of the canonical start site (uORFs) in the 5' untranslated regions (5'UTRs) of their mRNAs can respond to ISR pathway-driven upregulation (26). In the absence of cellular stress, translation initiation at these upstream starts suppresses translation at the canonical start codon, and inhibits synthesis of these stress mediators. Inhibition of ternary complex recycling by increased eIF2 $\alpha$  phosphorylation leads to increased read-through of the upstream starts, allowing increased translation initiation at the canonical start site (27, 28). In yeast, the transcription factor GCN4 is the key mediator of ISR pathway-driven transcriptional changes (29). In vertebrates, its functional equivalent is ATF4. Each of these stress-mediating mRNAs contains multiple uORFs that confer dramatic upregulation of these transcription factors in response to eIF2 $\alpha$  phosphorylation (30). Other uORF-containing transcripts that are translationally upregulated during ISR pathway activation include CHOP and CYP2B (31). Although uORFs are necessary for the translational upregulation component of stress-driven ISR pathway response, they do not appear to be sufficient to confer this response. Additional parameters important for uORF regulation (uORF number, frame, or position relative to canonical start) are poorly understood (26, 32, 33).

Cellular stress activates the ISR via one or more eIF2 $\alpha$  kinase that then phosphorylate eIF2 $\alpha$  (**Fig 1**). Mammals have a set of four eIF2 $\alpha$  kinases with similar kinase domains, but distinct regulatory domains that facilitate activation by specific inputs allowing each kinase to primarily respond to a distinct type of cellular stress (9, 34). These kinases include GCN2 (general control non-derepressible-2), PERK (PKR-like ER kinase), HRI (heme-regulated inhibitor), and PKR (protein kinase double-stranded RNA-dependent) (9).

The number of eIF2 $\alpha$  kinases and the range of cellular inputs activating the ISR expanded from yeast to mammals with one kinase represented in *S. cerevisiae* and four in mammals (**Fig 2**) (9, 24). GCN2 is represented throughout eukaryotes and is activated by uncharged transfer RNAs (tRNA). In yeast GCN2 responds to amino acid, glucose, or purine starvation, and high salinity (35). In mammals, GCN2 responds to amino acid and glucose starvation, but also UV irradiation and proteasome inhibition (10). PERK is an endoplasmic reticulum (ER) trans-membrane protein that has only been found in metazoans, which is activated by unfolded proteins in the ER, referred to as ER stress (36). The extent of conservation of HRI across eukaryotes is not known and its absence in traditional invertebrate models has led to the misinterpretation of HRI as a vertebrate specific eIF2 $\alpha$  kinase (37, 38). It has been found in mammals, some species of arthropods, and putatively in one species of fungi, but not *S. cerevisiae* (37-39). HRI senses low heme levels and also responds to heat shock and heavy metals (40, 41). Lastly, PKR has only been identified in vertebrates, where it senses double-stranded RNA responding to viral infection (9). ISR-mediated translation inhibition allows the cell to preserve energy and amino acids during amino acid deprivation, reduce the ER protein load to allow the ER to fold or dispose of misfolded proteins during ER stress, and couple the synthesis of globin to cellular heme levels particularly when heme levels are low (10, 37, 42). In addition, ISR induction increases transcription of a subset of genes that aid the cell in its recovery. Conservation of GCN2 across existing eukaryotes indicates that a GCN2-like kinase was present in the first eukaryotes that later diversified, but the evolution of these kinases is not well understood (11, 12).



The conservation of mammalian ISR gene targets has not been investigated in non-vertebrate metazoans. As a consequence of the shared phosphorylation of eIF2 $\alpha$  by multiple kinases in mammals, common ATF4 responsive genes are transcribed in response to diverse stresses such as amino acid limitation and ER stress (21, 43). ATF4 targets include genes important in cell recovery from stress or ultimately apoptosis if the stress is unresolved. Characteristic mammalian ATF4 targets include genes involved in: amino acid metabolism and transport including ASNS (asparagine synthetase) (44-46), SNAT2 (sodium-coupled neutral amino acid transporter 2) (47, 48), and xCT (cysteine/glutamate transporter) (49, 50); proteostasis and oxidative balance including HERP1 (homocysteine-inducible, ER stress-inducible, ubiquitin-like domain member 1) (51-54) and CHAC1 (glutathione-specific gamma-glutamylcyclotransferase 1) (55); and negative feedback regulation of the ISR including GADD34 (growth arrest and DNA damage-inducible protein) (56), CHOP (CCAAT-enhancer-binding protein homologous protein), and TRB3 (tribbles homolog 3) (57-59). In *S. cerevisiae*, GCN4 elicits transcriptional activation of at least 539 genes, primarily amino acid biosynthetic genes (35). Notably, ASNS is the best-characterized mammalian ISR induced target; however, yeast ASNS homologues are not ISR induced (35, 45, 60, 61). Characterization of the conservation of these targets in *Nematostella* could yield insight into the function and limits of ISR-induced changes to stress adaptation.

### **Using ISR pathway activation to monitor marine environmental and organismal health**

Marine environments are increasingly impacted by anthropogenic activities. Over 50% of the human population lives in coastal areas in the United States and coastal waters are especially impacted by human activity (62). The estuarine habitat of *Nematostella* constitutes a particularly ecologically and economically important ecosystem making accurate monitoring of habitat health important (62). Compounding the natural variability of these ecosystems, estuarine environments are highly impacted by human activities resulting in the introduction of toxic chemicals and alterations in fresh and saltwater inputs (63). Climate change is likewise predicted to cause increasing temperatures (64). *Nematostella* as an infaunal and largely sedentary animal is particularly exposed to accumulation of toxic chemicals in sediments and physical changes in its immediate environment, suggesting the use of *Nematostella* as a sentinel organism. In addition, estuaries exhibit natural variability in major environmental parameters including temperature, salinity, ultraviolet radiation, oxygen, and reactive oxygen species (65, 66) and as a result, estuarine species such as *Nematostella*, exhibit broader physiological tolerances than freshwater, marine, or terrestrial species (63). The biochemical basis for this broad tolerance in *Nematostella* is not fully understood. Elevated temperatures, changes in salinity, and heavy metals are only a subset of the stressors that can activate the ISR in eukaryotic cells that are known environmental stressors affecting marine organisms.

Environmental stressors are diverse, but often cause similar damage or alterations to the cell. Elevated temperature and hypertonic stress primarily induce protein misfolding in the cytoplasm, while heavy metal exposure induces protein misfolding in the ER. These stressors trigger ISR pathway activation in mammalian cells (67-71). Protein misfolding and subsequent ISR induction has been most studied in the ER. The reducing agent, dithiothreitol (DTT) is a well-established ER stress inducer that disrupts disulfide bonds resulting in misfolded ER proteins. In the experiments described below, I utilize DTT to mimic a range of environmental stressors. Treatment of *Nematostella* with DTT produces cellular ER stress and subsequent activation of the ISR pathway in this organism. This observation is the first demonstration of induction of the ISR in an animal from a basal lineage, as well as the first demonstration of an intact ER stress response in *Nematostella*.

Analysis of the *Nematostella* genome predicts conservation of the core mammalian ISR. Likewise, comparison of eIF2 $\alpha$  orthologues predicts a high degree of epitope conservation, which, in fact, has proved sufficient to allow the use of commercial mammalian antibodies for the detection of both

eIF2 $\alpha$  and p-eIF2 $\alpha$  proteins in these studies. By exposing *Nematostella* to a range of chemical and physical triggers, then assessing the phosphorylation state of eIF2 $\alpha$  and putative ISR transcriptional mediators, I have identified and begun to characterize multiple arms of the ISR in this organism. In these experiments, I use chemicals, (DTT and halofuginone (HF)), which mimic diverse environmental stress (oxidative stress and dietary restriction, respectively). I further show that exposure to conditions that mimic physical environmental stressors and pollution in estuaries, such as elevated temperature, increased salinity, and copper results in increased eIF2 $\alpha$  phosphorylation or ISR target gene transcription, thereby coupling ISR pathway activation to environmental stress in *Nematostella*.

## RESULTS

### **The *Nematostella* genome contains major components of the mammalian ISR**

To determine whether the major components of the ISR are present in the basal metazoan, *Nematostella*, I searched the *Nematostella* genome for the presence of orthologues to the mammalian kinases, the central transcription factor, ATF4, and downstream ISR transcriptional target genes. Human protein query sequences were blasted against the Joint Genome Institute (JGI) assembly of the *Nematostella* genome and were considered orthologues if reciprocal blasts against the human genome yielded positive matches. The nucleotide sequences of positive JGI hits were further blasted against the *Nematostella* transcriptome published by Tulin et al. 2013 to determine whether the JGI sequence represented the full transcript (72).

Three of the four mammalian eIF2 $\alpha$  kinases, GCN2, PERK, and HRI have orthologues in the *Nematostella* genome (**Fig 3, Fig S1**). Phylogenetic analysis of eIF2 $\alpha$  kinases from 13 eukaryotic genomes including *Nematostella*, yeast, and 11 additional metazoan species shows that the *Nematostella* kinases fall within expected kinase member clades, supporting these *Nematostella* kinases as true orthologues of mammalian GCN2, PERK, and HRI (**Fig S1**). Specific residues have been previously identified in mammalian eIF2 $\alpha$  kinases that distinguish eIF2 $\alpha$  kinases from other kinases (73). These residues are present in all three *Nematostella* kinases (data not shown), establishing these kinases as eIF2 $\alpha$  kinases. eIF2 $\alpha$  kinases contain similar kinase domains responsible for phosphorylating eIF2 $\alpha$  but have distinct regulatory regions that respond to distinct cellular signals (**Fig 3a**). Sequence identity between the kinase domains of each *Nematostella* eIF2 $\alpha$  kinase with its respective human orthologue is greater than between any of the other human eIF2 $\alpha$  kinase family members (**Fig 3b**).

To determine whether the characteristic domains of each of the mammalian kinases are present in the *Nematostella* kinases, the *Nematostella* kinases were searched for conserved domains using NCBI Conserved Domain Search. NvGCN2 contains the characteristic pseudo-kinase and histidinyI-tRNA domains found in yeast and mammalian GCN2 that are required for kinase activity and uncharged tRNA binding, respectively (74). In addition NvGCN2 contains the characteristic RWD domain at its N-terminal domain responsible for protein interaction (**Fig 3a**). NvPERK contains a luminal domain that is responsible for dimerization and sensing of unfolded proteins (75). A putative ER signal peptide or long stretch of hydrophobic residues is also present within the N-terminus (N-CFSSVVKFVAVMATFAN EFGNVSPN-C) (**Fig 3a**). The kinase domain of NvHRI is most similar to HRI; however the N-terminal portion of NvHRI (responsible for heme binding and detection in mammalian HRI) shares low sequence identity (26.0%) with mammalian HRI (**Fig 3a**). Only a subset of residues within this region may be required for heme binding. Residues important for heme binding in mammals including His75, His119, His120, and Cys409 are not conserved and there are no Cys-Pro motifs, which in mammals participate in heme binding (76, 77). Functional studies will be required to confirm if NvHRI responds to heme levels. Given the importance of heme and hemeproteins in cells (gas transport, chemical catalysis, and electron transfer) and the presence of HRI in a wide range of eukaryotic species it would not be surprising if NvHRI responds to heme levels. The presence of NvHRI confirms that HRI is not restricted to vertebrates but is likely represented throughout eukaryotes and has been lost in several lineages including *S.*

*cerevisiae*, *C. elegans* and *D. melanogaster*. Conservation of NvPERK demonstrates that PERK appeared evolutionarily as early as *Nematostella* on the metazoan evolutionary tree. PERK was already present in the cnidarian-bilaterian ancestor supporting PERK as a metazoan-specific kinase. The lethal phenotype of PERK null mice and the appearance of endogenous ER stress with metazoan development and cellular differentiation suggest PERK may have been important in mediation of endogenous ER stress during evolution (75, 78). Within all three *Nematostella* eIF2 $\alpha$  kinases, threonine residues responsible for kinase autophosphorylation in mammals are conserved. This demonstrates that *Nematostella* contains orthologues to 3 of the four mammalian kinases.

### **ATF4 and ISR transcriptional targets conserved in *Nematostella***

The transcription factor, ATF4, is currently considered the primary mediator of early ISR-induced transcriptional changes (79). An ATF4 bZIP domain was recently identified in the *Nematostella* genome; however, this study did not identify the full coding sequence and this gene is not included in the JGI predicted gene database (80). This explains why a blast search of the *Nematostella* genome with the human protein sequence yields no predicted protein hits within JGI. The nucleotide sequence of the NvATF4 bZIP domain is however present in the JGI genome and a full transcript encompassing this NvATF4 bZIP domain can be found in the *Nematostella* transcriptome published by Tulin et al. 2013 (72). This transcript designated NvATF4 shares only 22.9% identity with its human counterpart (**Fig 1b**), but shares 44.2% identity within the bZIP domain. Phylogenetic analysis of bZIP domains from bZIP transcription factors across 13 eukaryotic species shows that NvATF4 groups with other metazoan ATF4 orthologues supporting NvATF4 as an orthologue of mammalian ATF4 (**Fig S2**). Blast searches of three other cnidarian species (the hydroid, *H. vulgaris*; the coral, *A. digitifera*; and the anemone, *A. elegantissima*) with full-length NvATF4 yield orthologues in these additional cnidarians, which share 20.3, 28.9, and 39.9% identity with *Nematostella*. Comparison of the bZIP domains of these orthologues also share higher identity than comparison between the full-length sequences (49.6; 59; 69.2%) suggesting NvATF4 orthologues are found throughout cnidarians. Additional reciprocal blast searches of two species from two of the earliest branching metazoan lineages: Trichoplax (*T. adhaerens*) and sponge (*A. queenslandica*), with full-length human ATF4 yields ATF4 orthologues in both these species supporting the presence of ATF4 across metazoans.

A key characteristic of ATF4 and GCN4, that is essential for ISR translational regulation is the presence of uORFs in the 5'UTR of the transcript (27). Analysis of the NvATF4 transcript shows the presence of four uORFs in the 5' UTR that are out of frame with the main ORF. A characteristic overlapping uORF of human ATF4 facilitates inhibition of ATF4 translation during non-stress conditions by allowing the ribosome to pass by the primary start site. Both the presence of NvATF4 and uORFs within the NvATF4 5'UTR suggests that transcriptional changes by the ISR are mediated similarly to mammals via upregulation of NvATF4 translation.

In mammals, ATF4 expression induces transcription of characteristic genes including both effectors of stress adaptation and negative feedback regulators of the ISR. To determine if these genes are present in *Nematostella*, human amino acid sequences were used in blast searches of the *Nematostella* genome and hits that yielded positive reciprocal blast matches either to the query itself or a closely related gene within the same protein family were considered orthologues or homologues, respectively. In mammals, the ISR is negatively regulated by two PP1 targeting proteins or eIF2 $\alpha$  phosphatases: GADD34 and CReP (56, 81). The first invertebrate eIF2 $\alpha$  phosphatase was recently identified in *D. melanogaster* (82). To identify a *Nematostella* eIF2 $\alpha$  phosphatase homologue, the *Nematostella* genome was searched as described but with the *D. melanogaster* phosphatase as query because the human query yielded no results. *Nematostella* contains one eIF2 $\alpha$  phosphatase designated NvEPP, which contains a PP1 binding KVxF motif in the C-terminal region like mammalian and fly GADD34. No other invertebrate eIF2 $\alpha$  phosphatases have been identified.

In addition to ISR regulatory genes, ATF4 expression induces genes involved in amino acid biosynthesis and transport, ER homeostasis, and oxidative balance. Mammalian ISR targets with identified *Nematostella* homologues are shown in **Fig 3c**. *Nematostella* contains orthologues of ASNS, HERP1, CHAC1, TRB3 and putative homologues or related proteins of mammalian SNAT2 and xCT (**Fig 3c**). NvASNS and NvHERP share the highest sequence identity with their human orthologues. ASNS is the best-studied ISR target, is highly conserved in *Nematostella* as expected, and shares 58% identity with human ASNS. In addition, within 2000 base pairs of the ASNS translational start site are 5 putative ATF4 binding sites, supporting NvATF4 regulation of NvASNS. Other *Nematostella* ISR target homologues also contain at least 2 putative ATF4 binding sites. This is the most basal metazoan lineage in which conservation of ISR target genes have been examined, and suggests that many mammalian ISR targets are highly conserved in this basal metazoan.

### **Phosphorylation of eIF2 $\alpha$ and putative ISR target genes are induced by the reducing agent, DTT, as in mammalian cells**

To test whether the ISR is induced similarly to the ISR in mammals, adult *Nematostella* were exposed to dithiothreitol (DTT). eIF2 $\alpha$  phosphorylation (peIF2 $\alpha$ ) and induction of putative ISR target genes were assayed by immunoblot and qPCR. DTT disrupts protein folding by disrupting disulfide bonds, causing ER stress and inducing the ISR. Phosphorylation of eIF2 $\alpha$  in *Nematostella* increased significantly following exposure to DTT for four hours (**Fig 4a left**). RNA levels of ISR targets were also induced significantly (**Fig 4a right**). ASNS showed the greatest fold increase. This induction is much greater than ASNS induction by ISR inducers in mammalian cells. HERP and TRB showed the greatest induction among the remaining target genes. The timecourse of peIF2 $\alpha$  and ISR target gene transcription was similar to that of mammalian cells. In mammalian cells peIF2 $\alpha$  is induced within minutes and declines after 4 to 6 hours (45), while ASNS transcription subsequently increases and starts to decline after 4-6 hours (4). This delay of transcriptional induction compared to peIF2 $\alpha$  is consistent with regulation of ASNS by peIF2 $\alpha$ . In *Nematostella* peIF2 $\alpha$  also occurred within 15 minutes of DTT exposure (**Fig 4b left**) while no ASNS transcription was observed after a 15 minute DTT exposure (data not shown). Transcription of ASNS and HERP increased within 2 hours of DTT exposure (**Fig 4b right**) and 2 hours after a 15-minute exposure to DTT (data not shown). Both phosphorylation and ASNS levels were significantly induced by 1 and 4 mM DTT (**Fig 4c**), which is within the range used to induce the ISR in mammalian cells (83). DTT induction of ASNS and HERP was also seen in juveniles (data not shown). The strong, rapid induction of both conserved ISR transcriptional responses and of eIF2 $\alpha$  phosphorylation by an ER stressor in *Nematostella* provides direct functional evidence that the ER stress component of the ISR is conserved in *Nematostella*, and responds to an ER stressor over a dose range and time course similar to that seen in mammalian cells.

To directly test whether PERK mediates eIF2 $\alpha$  phosphorylation by DTT in *Nematostella*, adults were exposed to the mammalian PERK inhibitor, GSK2606414, followed by exposure to DTT, and peIF2 $\alpha$  levels assayed. Decreased eIF2 $\alpha$  phosphorylation in 293T cells with PERK inhibitor followed by DTT exposure confirmed inhibition of mammalian PERK (**Fig 5a**). In *Nematostella* only the highest dose (100  $\mu$ M) of inhibitor decreased phosphorylation (**Fig 5b**). While this could be due to inhibition of NvPERK, the high dose (1000x the mammalian dose) required makes it difficult to conclude that the inhibitor effects are specific to this kinase. Whether the low efficacy of GSK2606414 in *Nematostella* is due to poor uptake or penetrance, divergence between the NvPERK and mammalian PERK kinase domains, or other factors, is not clear.

In vertebrates and *D. melanogaster* peIF2 $\alpha$  is enriched in tissues that secrete large amounts of proteins, and therefore have an elevated protein-folding load in the ER (78). To examine localization of peIF2 $\alpha$  in *Nematostella*, immunohistochemistry was performed with peIF2 $\alpha$  antibody in juveniles treated with or

without DTT (**Fig S3**). DTT-stress stimulated peIF2 $\alpha$  staining is detectable and is localized towards the oral end of the animal (**Fig S3b, h**). Localization of peIF2 $\alpha$  in unstressed animals was not clear, suggesting that there is little endogenous peIF2 $\alpha$  in these specimens or it is significantly weaker than that of DTT exposed animals and was not detected here (**Fig S3a**). Localization of peIF2 $\alpha$  in juvenile *Nematostella* shows no obvious pattern at the level of resolution provided by the experimental setup. Comparison of peIF2 $\alpha$ -specific FITC and nuclear DAPI staining does suggest that NvpeIF2 $\alpha$  is localized to the cytoplasm, as expected. Sectioned animals and confocal microscopy, along with additional antibody staining to identify specific cell types will be necessary to determine whether peIF2 $\alpha$  is enriched in certain parts or cells of this animal. A *Nematostella* specific antibody could be raised; however, due to the high conservation of the serine phosphorylation site and surrounding sequence it is unlikely that the *Nematostella* specific antibody would be advantageous. These results replicate the previous results of increased peIF2 $\alpha$  phosphorylation after exposure to the ER stress inducer, DTT.

### **Halofuginone, a chemical inducer of the GCN2 arm of the mammalian ISR, induces the ISR in juvenile *Nematostella***

The GCN2 arm of the ISR is often referred to as the amino acid response (AAR). Halofuginone (HF) is an AAR pathway activator that specifically binds and inhibits the enzyme, prolyl tRNA synthetase (EPRS) leading to the production of unbound tRNAs and the phosphorylation of GCN2 (84). EPRS is conserved in *Nematostella* suggesting activation of NvGCN2 by HF is likely (85). To determine whether HF activates the ISR in *Nematostella*, adults and juveniles were exposed to HF, and assayed for eIF2 $\alpha$  phosphorylation and induction of the transcriptional readouts most highly induced following DTT exposure: ASNS and HERP. HF significantly increased eIF2 $\alpha$  phosphorylation and induction of ASNS in juvenile *Nematostella* (**Fig 6**); however, neither eIF2 $\alpha$  phosphorylation nor ASNS and HERP transcription were induced in adults (data not shown). This could indicate a difference in the physiology of juveniles compared to adults or HF is simply able to permeate the smaller juveniles more efficiently. The significant induction of peIF2 $\alpha$  and ASNS support the use of HF to activate the AAR via GCN2 in *Nematostella*, and possibly as a method to mimic amino acid starvation.

### **Exposure to the environmental stressors: temperature, salinity, and copper activates the ISR in *Nematostella***

To evaluate the relevance of the ISR to environmental stressors that threaten marine and human health, induction of peIF2 $\alpha$  and the transcriptional targets: NvASNS and NvHERP, shown to increase after exposure to two mammalian ISR inducers were assayed subsequent to exposure to low and high temperatures, low and high salinities, and copper. In yeast and mammals, heat shock results in eIF2 $\alpha$  phosphorylation (38, 39). Temperatures collected in a pool in Sippewissett Marsh, MA showed that temperatures varied between 1.3 and 43.7°C and therefore *Nematostella* must tolerate wide variations in temperature (14). In *Nematostella* phosphorylation of eIF2 $\alpha$  increased significantly following a 30 minute exposure to 37°C (**Fig 7a**). Longer exposure of adult *Nematostella* to 37°C (1 hour or more) was lethal to the animals, and the lack of peIF2 $\alpha$  induction at these times may reflect cell death. It is possible that as infaunal organisms, animals in the field do not actually experience such high temperatures. In addition, animals in the field likely do not experience such abrupt changes in temperature, perhaps allowing adaptation to subsequent higher temperatures not replicated in these experiments. As observed for DTT, eIF2 $\alpha$  phosphorylation was induced by heat shock within 10 minutes of exposure (**Fig 7b**). Exposure to 4°C did not induce eIF2 $\alpha$  phosphorylation (**Fig 7c**). NvASNS transcription was stimulated 3.5 hours after a 30 minute exposure to 37°C (**Fig 7d**). In addition, transcription of the ER chaperone, BiP (grp78), a member of the Hsp70 family and a major effector of ER stress adaptation, was induced by brief heat stress (**Fig 7d**).

Because *Nematostella* inhabit pools that undergo large changes in salinity, these animals must adapt to both hypertonic and hypotonic stresses. The AAR can mediate adaptation to hypertonic stress in both *C. elegans* and yeast (67, 86, 87), and therefore I examined the effects of salinity changes on *Nematostella*. *Nematostella* have been documented to survive salinities ranging from 2 – 54 parts per thousand (ppt) and are raised in the lab at 15 ppt (88). Exposure to 0 or 50 ppt artificial seawater for four hours significantly induced NvASNS expression at 50 ppt but not 0 ppt (**Fig 7e**) consistent with the possibility that hypertonic stress induces the ISR possibly via NvGCN2 in *Nematostella*. Further experiments will be required to determine that peIF2 $\alpha$  is induced and whether NvASNS induction by increased salinity is due specifically to NvGCN2 and ISR activation.

Copper is a common marine pollutant, and copper and other metals have been shown to induce the ISR. For that reason I examined the effects of copper exposure on NvASNS and NvHERP expression (40, 70, 89). Trace metals are particularly prevalent in estuaries, and accumulate to levels orders of magnitude higher in coastal waters than in the open ocean (90). Metal ions can directly interact with cellular proteins, indirectly interact through metal-induced oxidative damage of proteins, and can affect the function of metal requiring proteins and enzymes leading to oxidative stress (90). Copper is a commonly used test example for metal toxicity and has been shown to induce ER stress (91-94). Metal toxicity is generally tested using sublethal concentrations, and in vertebrate cells copper concentrations ranging from 1-50  $\mu$ M are typically used (89, 92, 93). *Nematostella* exposed to 20  $\mu$ M copper significantly induced NvASNS expression but not NvHERP (**Fig 7f**), consistent with ISR induction, but additional experiments will be required to determine whether copper induces NvpeIF2 $\alpha$  and the specificity of induction to the ISR. These results are consistent with activation of the ISR by heat stress, hypertonic stress, and copper exposure indicating a role for the ISR in environmental stress response in *Nematostella*.

## DISCUSSION

These results establish that the major components of the mammalian ISR are present in *Nematostella vectensis* and provide evidence for ISR pathway activation by chemicals and environmental stresses. The expansion of inducers, effectors, and functions of the ISR seen in mammals compared to those present in yeast were most likely present in the last common ancestor between cnidarians and bilaterians. The conservation of NvPERK confirms that PERK was present in the cnidarian-bilaterian ancestor, backing PERK as a metazoan-specific kinase. In addition, the presence of NvHRI confirms that HRI is not restricted to vertebrates but is likely represented throughout eukaryotes and has been lost in several lineages. The *Nematostella* genome contains an ATF4 orthologue with uORFs suggesting similar function to mammalian ATF4 and that the central function of ATF4 in the transcriptional response of the ISR is conserved as far back as cnidaria. Finally several targets of the mammalian transcriptional response are conserved and are induced by chemical and environmental stresses known to induce the ISR in other species. ASNS in particular is highly induced by both the chemical and environmental stresses used. In particular, DTT results in dramatic induction of ASNS and strong induction of peIF2 $\alpha$ . ER stress can be caused by a variety of environmental stressors underscoring the likely central role of the ISR in stress response in *Nematostella* and other marine organisms. The high conservation of the central core of this pathway in basal metazoa demonstrated by this chapter suggest the application of these tools for the study of the ISR and stress response in other non-model marine metazoans.

The ecology of *Nematostella* and the conservation of many components of the mammalian ISR underscore the promise of *Nematostella* as an indicator species of environmental health. Transgenic mice, flies, and zebrafish carrying ISR reporters have been successful in indicating physiological stress due to endogenous and external sources (89, 95, 96). A recent study took advantage of a transgenic zebrafish line harboring a human upstream open reading frame of an ISR translational target gene fused with GFP. The reporter reliably reflected physiological stress induced by exposure to heavy metals and endocrine disrupting chemicals. Intriguingly, different tissue-specific GFP expression patterns were

observed in response to different chemical stressors. Animals exposed to field-collected water samples were able to correctly detect various pollutants (89). The advantage of this sort of transgenic model is that it reflects the physiological stress of the organism without respect to a specific stressor allowing detection of unexpected hazardous chemicals, stresses, and the real-life reality of multiple stress exposure, while traditional chemical analysis and transcriptional targets of known stressors can provide specificity. The ideal ecology of *Nematostella* as an estuarine, sedentary, yet stress-resistant animal could provide a perfect system as a similar transgenic model for estuarine health. The work described here, which identifies and characterizes intact and functional ER stress and AAR pathway arms of the ISR in *Nematostella*, lays the groundwork for the development of *Nematostella* into a sentinel organism for an essential and highly threatened marine ecosystem.

Proteostasis, or the maintenance of the proper function of cellular proteins, is disrupted by multiple stressors and causes ER stress and therefor likely activates PERK. Recent studies in *C. elegans*; however, implicate GCN2, and thus the AAR pathway arm of the ISR, in proteostasis (67, 86). Since maintenance of proteostasis is central to adaptation and the potential toxicity of environmental stressors, understanding the regulation of proteostasis should be central to understanding environmental stress toxicity and the capacity for organismal stress adaptation. Although reduction in protein synthesis is widely observed in response to environmental stress, the relationship between reductions in protein synthesis and improved stress resistance is not understood (86). Inhibition of translation initiation in *C. elegans* during hypertonic stress occurs via activation of GCN2, and translation inhibition is required to reduce protein damage during hypertonic stress (67, 86). Induction of ASNS by hypertonic stress in *Nematostella* suggests that the ISR may play a similar protective role during hypertonic exposure. Whether the *Nematostella* hypertonic response occurs through GCN2, as in *C. elegans* and the role of this response in proteostasis will be interesting areas for further investigation.

GCN2 and the AAR are responsible for adaptation to nutrient limitation and contribute to the beneficial effects of lifespan extension and prevention of age-related diseases observed in model metazoan species in response to caloric restriction (84). *Nematostella* are able to survive extended periods of time without food, during which they shrink, but quickly regain size upon feeding. Phosphorylation of eIF2 $\alpha$  is elevated in highly stress resistant states of other animals such as hibernating mammals, implicating ISR involvement in these survival states (97). AAR pathway activation following amino acid deprivation in mammals is known to induce autophagy, suggesting the ISR may play a significant role in mediating survival during hibernation and extended states of caloric restriction (23). The work presented here provides a basis for future studies of these phenomena.

Through the exposure of *Nematostella* to a range of chemical and physical triggers, I've demonstrated the conservation of and characterized the activation of multiple arms of the ISR pathway in this organism. The work presented here demonstrates the conservation of the mammalian core of the ISR in a basal metazoan and validates assays that can be used to follow known components of ISR activation in a range of non-model metazoans, where experimental tools for functional protein studies are limiting. Further work in this thesis will explore new pathways, distinct from the core ISR, that mediate responses to amino acid limitation, and examine whether these additional stress responses are also shared across metazoans.

## METHODS

### Identification of ISR genes in *Nematostella*

ISR target genes were identified in the *Nematostella* genome through BLASTp searches of the Joint Genome Institute (JGI) *Nematostella vectensis* assembly using the human amino acid ISR gene sequences whose accession numbers are given in **Table S1**. The JGI nucleotide sequences were blasted against the *Nematostella* developmental transcriptome published by Tulin et al. 2013 using BLASTn to assess whether the JGI sequence was complete (72). In cases where multiple sequences matched, sequences were assembled *in silico* to produce the most complete sequence. The identifiers of the sequences used in this study are given in **Table S1**.

### Phylogenetic Analyses of eIF2 $\alpha$ Kinases and bZIP Transcription Factors

To determine evolutionary relationships of Nv eIF2 $\alpha$  kinases with other metazoan eIF $\alpha$  kinases, eIF2 $\alpha$  kinase homologues from additional species were obtained by using amino acid sequences of the four human eIF2 $\alpha$  kinases (accession numbers supplied in **Table S1**) to search *Saccharomyces cerevisiae*, *Amphimedon queenslandica*, *Acropora digitifera*, *Nematostella vectensis*, *Hydra vulgaris*, *Caenorhabditis elegans*, *Drosophila melanogaster*, *Ciona intestinalis*, *Strongylocentrotus purpuratus*, *Branchiostoma floridae*, *Petromyzon marinus*, *Danio rerio*, and *Homo sapiens* genomes using BLASTp. *S. cerevisiae*, *C. elegans*, *C. intestinalis*, *D. melanogaster*, *D. rerio*, and *P. marinus* genomes were searched through ENSEMBL, the *A. digitifera* genome was searched through the OIST Marine Genomics Genome Browser, the *A. queenslandica* genome was searched through ENSEMBL METAZOA, and *Branchiostoma floridae*, *Strongylocentrotus purpuratus*, and *Hydra magnipapillata* were searched through NCBI (98, 99). Sequences used for subsequent analysis were those with the highest scores, lowest e-values, and that yielded reciprocal best BLAST hits. Protein accession numbers of protein sequences are given in **Table S1**. Sequences were aligned with Muscle 3.6 and manually edited in Geneious 5.6.6 (100, 101). ProtTest was used to determine the best model of protein evolution (102). Maximum likelihood analysis was run using RAxML v.8.0.0 (103) with the following model and parameters: LG+I+G+F. One thousand rapid bootstraps were performed. Figtree v.1.4.0 was used for tree visualization (104).

To determine the evolutionary relationships of Nv bZIP transcription factors with other eukaryotic bZIP transcription factors, bZIP transcription factors from *Saccharomyces cerevisiae*, *Monosiga brevicollis*, *Nematostella vectensis*, *Caenorhabditis elegans*, *Drosophila melanogaster*, *Ciona intestinalis*, and *Homo sapiens* from Amoutzias et al. 2007, Fassler et al. 2002, Reinke et al. 2013, Vinson et al. 2002, and Yamada et al. 2003 were used to construct a maximum likelihood tree. bZIP domain sequences for all taxa were aligned with Muscle 3.6 and manually edited in Geneious 5.6.6 (100, 101). The subsequent maximum likelihood analysis was performed as described above for the eIF2 $\alpha$  kinases with the following model and parameters: LG+G.

### *Nematostella* Culture

*Nematostella vectensis* were cultured under standard conditions (0.45  $\mu$ m filtered seawater (for experiments conducted at WHOI) or artificial seawater (Instant Ocean) (for experiments conducted at HSDM) diluted to 13 ppt, four weekly brine shrimp feedings, and weekly water changes) similar to previously described conditions (105). Spawning was performed according to (106) with additional mussel ovary feeding the day prior to spawning.

### Mammalian Cell Culture

293T (ATCC) were cultured at 5% CO<sub>2</sub> at 37°C in DMEM supplemented with 10% fetal bovine serum, L-glutamine, non-essential amino acids (NEAA), and antibiotics.



### ***Nematostella* Chemical and Environmental Stress Exposures**

Adult *Nematostella* starved for three to four days were transferred to 24-well tissue culture plates with 13 ppt seawater. One animal was placed in each well. After transfer, plated animals were incubated at room temperature overnight. For chemical and salinity exposures, water was replaced with 1 mL per well of 13 ppt seawater containing the appropriate concentration of chemical (DTT, HF (Whitman lab), PERK inhibitor (GSK2606414, Selleckchem), copper sulfate (Sigma), or sea salt (Instant Ocean)). For temperature exposures individual animals were placed in 1.7 mL centrifuge tubes with 1 mL of 13 ppt seawater and placed on a heat block at the appropriate temperature.

### **Protein Extraction, SDS PAGE, and Immunoblotting**

For protein extraction, animals were placed on ice and rinsed 2x with 13 ppt seawater. For experiments performed at WHOI, animals were homogenized with a 1 mL dounce homogenizer (Wheaton) in 750  $\mu$ L high salt RIPA buffer (400 mM NaCl, 1% NP-40, 0.5% DOC, 0.1% SDS, 50 mM Tris pH 7.5, 1 mM EDTA with protease and phosphatase inhibitors: 20 nM Calyculin A, 60 nM Okadaic Acid, 1 mM PMSF, 1 mM NEM, and 1X Complete protease inhibitor cocktail (Roche)). Samples were incubated for 30 minutes at 4°C on a rotisserie tube rotator. After centrifugation (4°C, 20 minutes, 16,000 G), the supernatant was collected and the protein content quantified with the Coomassie Plus (Bradford) protein assay kit (Thermo Scientific). For experiments performed at HSDM, animals were homogenized in 500  $\mu$ L high salt RIPA at 4°C in Red Bullet Bead Lysis kit centrifuge tubes (Next Advance) for 15 minutes at level 8 in a bullet blender homogenizer (Next Advance). After centrifugation (4°C, 20 minutes, 16,000 G), the supernatant was collected and the protein content quantified with the Pierce BCA protein assay kit (Thermo Scientific). For both methods one animal was homogenized per sample.

Protein samples were diluted in 3X Laemmli buffer (160 mM Tris pH 6.8, 6% SDS, 20% glycerol, 0.01% bromophenol blue, 0.04%  $\beta$ -Mercaptoethanol), boiled for 5 minutes, and electrophoresed through pre-cast Bolt 4-12 % Bis-Tris Plus gels (Invitrogen). The amount of total protein loaded per well was standardized across each experiment. Proteins were transferred to nitrocellulose membrane (GE Healthcare), blocked with 5% milk and blotted with the following antibodies: eIF2 $\alpha$  (pS51), eIF2 $\alpha$ ,  $\beta$ -tubulin, actin (Sigma) (all from Cell Signaling Technology except actin). HRP-conjugated secondary antibodies were purchased from Cell Signaling Technology and blots were imaged with the Syngene PXi imager. Bands were quantified with ImageJ 1.48 (Wayne Rasband, National Institutes of Health).

### **RNA Extraction and Quantitative Real-Time PCR**

For RNA extraction, animals were placed on ice and rinsed 2x with 13 ppt seawater. For experiments performed at WHOI, animals were homogenized with a Teflon homogenizer in 1 mL PureZOL (Bio-Rad) and purified with the Aurum Total RNA Fatty and Fibrous Tissue kit (Bio-Rad) with 3-6 animals per sample. RNA extractions performed at HSDM were extracted with 1 mL TRIZOL (Ambion), homogenized in a bullet blender homogenizer in Red Kit RNase-free tubes for 10 minutes at level 10, and purified with the Direct-Zol RNA MiniPrep kit (Zymo Research) with 1-2 animals per sample. RNA yield and purity were quantified using an ND-1000 spectrophotometer and cDNA was synthesized using the iScript cDNA Synthesis Kit (Bio-Rad) at WHOI or the ABI High Capacity cDNA Reverse Transcription Kit (Ambion) at HSDM. Primer pairs were designed using Primer3 and primer sequences are available in **Table S2**.

Quantitative real-time PCR was performed using FastStart Universal SYBR Green Master Mix (ROX) (Roche) on the Step One Plus Machine (Thermo Fisher Scientific). The conditions were as follows: 95°C for 10 minutes, 40 cycles of 95°C for 10 seconds followed by 60°C for 1 minute. The qPCR efficiencies [ $E_{\text{gene}} = (10^{(-1/\text{slope})} - 1) * 100$ ] were determined by serial dilutions of pooled cDNA samples using cycle threshold (Ct) numbers for each gene analyzed. Efficiencies for all primer pairs were between 90 and 110%. Specificities of all PCR reactions were confirmed by melt curve analyses. Relative gene

expression changes were measured using the comparative  $C_T$  method,  $X=2^{-\Delta\Delta C_T}$ . Target gene expression was normalized with 18S as the endogenous control.

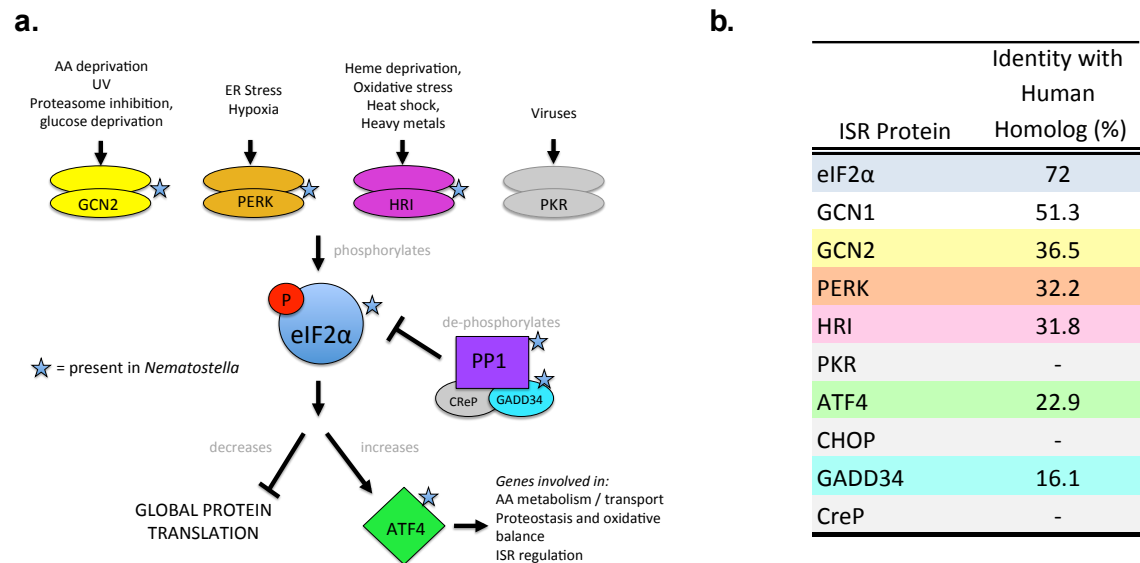
### **Immunohistochemistry**

Juvenile *Nematostella* (1 month old, unfed) were transferred to a 24-well plate (roughly 50 juveniles per well). Prior to fixation several drops of 7%  $MgCl_2$  were added to prevent anemone contraction. Anemones were allowed to sit for 10 minutes before fixation. Immunohistochemistry was performed as described in (107) with the following modifications: 1 mL ice-cold fixation solution was added directly to wells (13 ppt artificial seawater, 4% formaldehyde, 50 mM MOPS, pH 7.4, 2 mM  $\beta$ -glycerophosphate, 1 mM NEM) and incubated for 20 minutes before transfer to 1.7 mL centrifuge tubes. Samples were incubated for 1 hour at room temperature on rocker in fixation solution and then transferred immediately into MeOH. Animals were rehydrated in a series of 5 washes of PBS/MeOH (30, 50, 70, 90, 100% PBS in MeOH). Specimens were viewed and imaged with the Keyence BZ-X710 all-in-one fluorescence microscope.

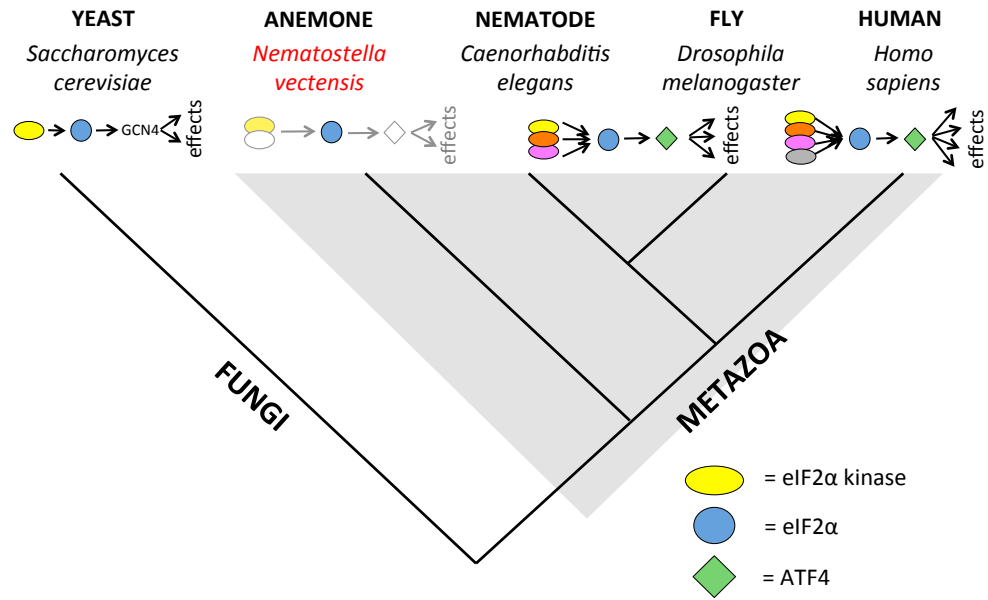
### **Statistics**

Data were analyzed using GraphPad PRISM, version 6. A probability value of 0.05 was considered statistically significant.

FIGURES AND TABLES

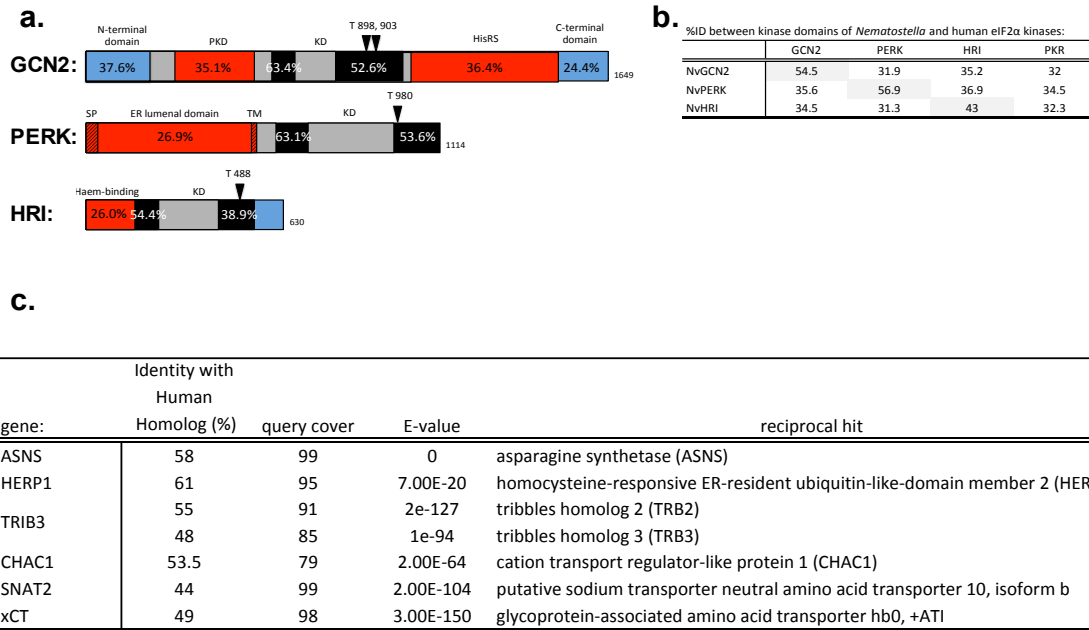


**Figure 1 | Cartoon representing basic components and features of the mammalian ISR and their presence in the *Nematostella* genome. (a)** eIF2α kinases are triggered by stress and phosphorylate eIF2α (indicated by red circle). eIF2α phosphorylation results in decreased levels of global translation and increased protein expression of ATF4, which further up regulates transcription of additional genes. ATF4 also up regulates the eIF2α-specific PP1 targeting protein, GADD34, which dephosphorylates eIF2α. Mammals also have another constitutively expressed eIF2α- specific PP1 targeting protein, CreP. Proteins with a blue star indicate orthologues that are found in the *Nematostella* genome (identification described later). **(b)** Table with ISR proteins in cartoon, indicating human and *Nematostella* accession numbers, as well as % identity between human and *Nematostella* orthologues, identification of which is described in the results section.

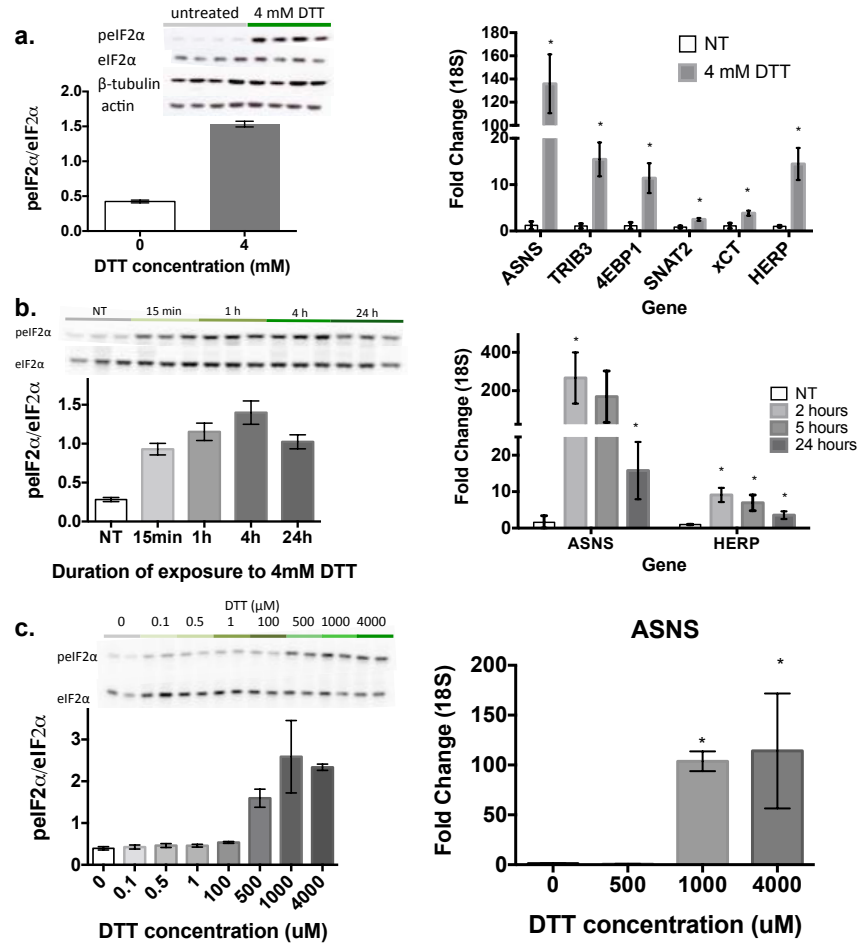


**Figure 2 | Cartoon schematic showing expansion of the inducers and effectors of the ISR from yeast to humans portraying the general evolutionary relationships between the model species depicted.**

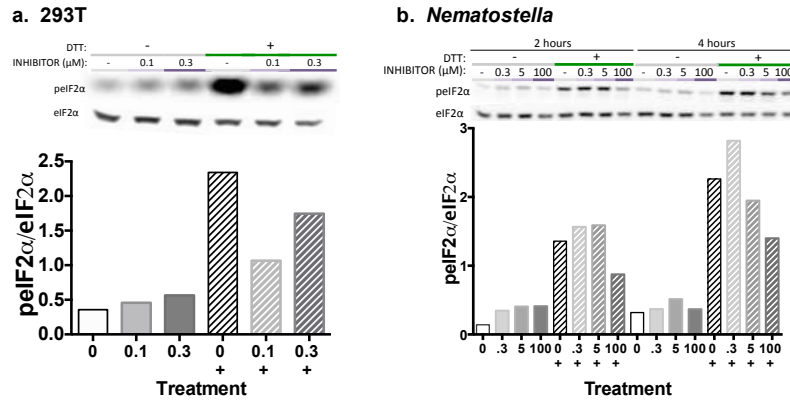
Figure indicates basic components of the ISR present in eukaryotic models. Below each species is a simplified pathway diagram to show which ISR components and effects are known to be present. Colors represent the respective proteins from the ISR pathway cartoon in figure one. This study characterizes ISR components present in *Nematostella*. *Nematostella* components are depicted in gray to emphasize that the ISR components in *Nematostella* have not been characterized. Colored *Nematostella* components indicate proteins predicted to be present based on conservation across eukaryotes.



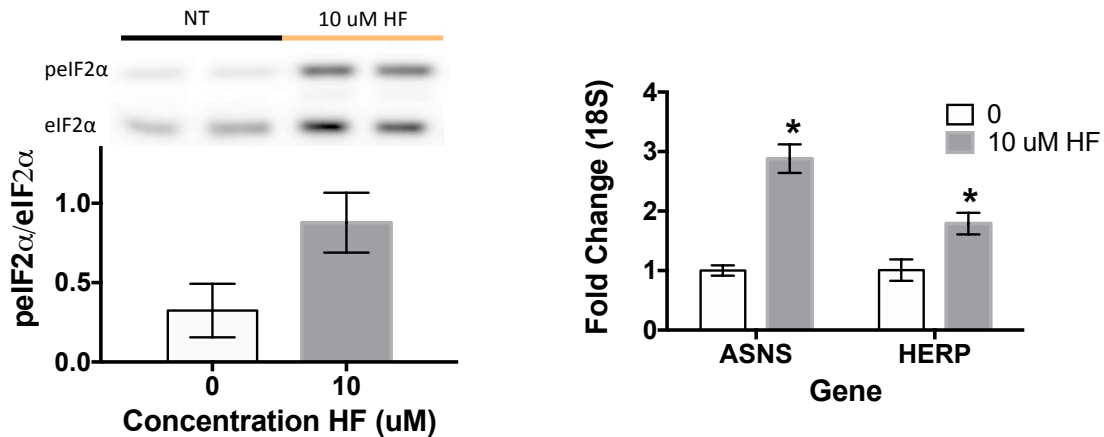
**Figure 3 | The three eIF2α kinases: GCN2, PERK, HRI and several ATF4 target genes are conserved in the *Nematostella* genome. (a)** Cartoon depicting 3 kinases conserved in *Nematostella* with characteristic domains labeled. Percentages indicate residue identity between human and *Nematostella* homologues within that domain. Black triangles indicate conserved threonine phosphorylation sites involved in autophosphorylation/activation of the respective kinase. KD, kinase domain; PKD, pseudo-kinase domain; HisRS, domain similar to histidyl-tRNA synthetase. Two black rectangles in each kinase represent the two lobes of the kinase domain with respective identities within each lobe indicated. **(b)** Table indicating percent identity of residues between kinase domains of human and *Nematostella* kinases. **(c)** Table of orthologues and homologues of mammalian ISR target genes identified in *Nematostella*.



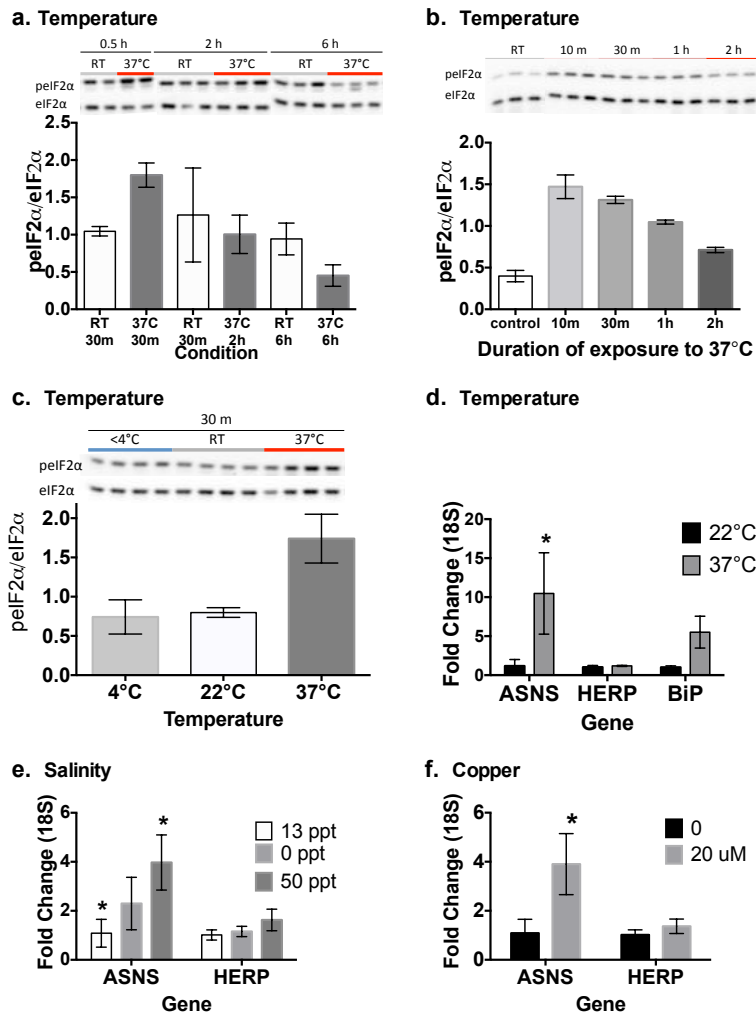
**Figure 4 | DTT increases phosphorylation of eIF2α and transcription of characteristic ISR responsive genes in *Nematostella*.** (a) (left panel) Adult *Nematostella* were treated with 4 mM DTT for 5 hours and assayed by immunoblot for total eIF2α and eIF2α phosphorylated on Ser51 using a commercial phospho-specific antibody. β-tubulin and cytoplasmic actin are shown as loading controls. (right panel) mRNA expression of characteristic ISR responsive genes indicated in figure from adults treated for 4 hours. (b) (left panel) Phosphorylated eIF2α and total eIF2α were measured by immunoblot in adults after 15 minutes, 1, 4, and 24 hours of 4 mM DTT treatment. (right panel) mRNA expression of ASNS and HERP after 2, 5, and 24 hours of 4 mM DTT treatment. (c) (left panel) Phosphorylated eIF2α and total eIF2α were measured by immunoblot in adults after 4 hours treatment with the indicated DTT concentrations. (right panel) mRNA expression of ASNS with the indicated DTT concentrations. (a-c) Densitometry data are displayed below the blot and presented as the mean ratio of phosphorylated eIF2α/total eIF2α ± s.d. from triplicate biological replicates. mRNA expression was normalized to expression of 18S mRNA and is represented as the mean of biological triplicates ± s.d. Data in (a) are representative of three separate experiments and data in (b) and (c) are one experiment. Confidence intervals (*P* value) for the effect of DTT versus untreated were determined using a one tailed Student's *t*-test. NT, nontreated; m, minutes; h, hours. All blots (except (c)) were also probed for actin to confirm equal loading and that total eIF2α levels did not change with treatment. (not shown).



**Figure 5 | PERK inhibitor treatment slightly decreases phosphorylation of eIF2α following DTT treatment in *Nematostella* but not as strongly as in the mammalian 293T cell line.** (a) 293T were treated for 2 hours with the indicated concentration of inhibitor in the presence or absence of 4 mM DTT for the second hour. Total eIF2α and phosphorylated eIF2α were assayed by immunoblot. (b) Adult *Nematostella* were treated for 2 or 4 hours with the indicated concentration of inhibitor in the presence or absence of 4 mM DTT for the last 15 minutes. Total eIF2α and phosphorylated eIF2α were assayed by immunoblot. (a-b) Densitometry data are displayed below the blot and data are representative of one experiment.



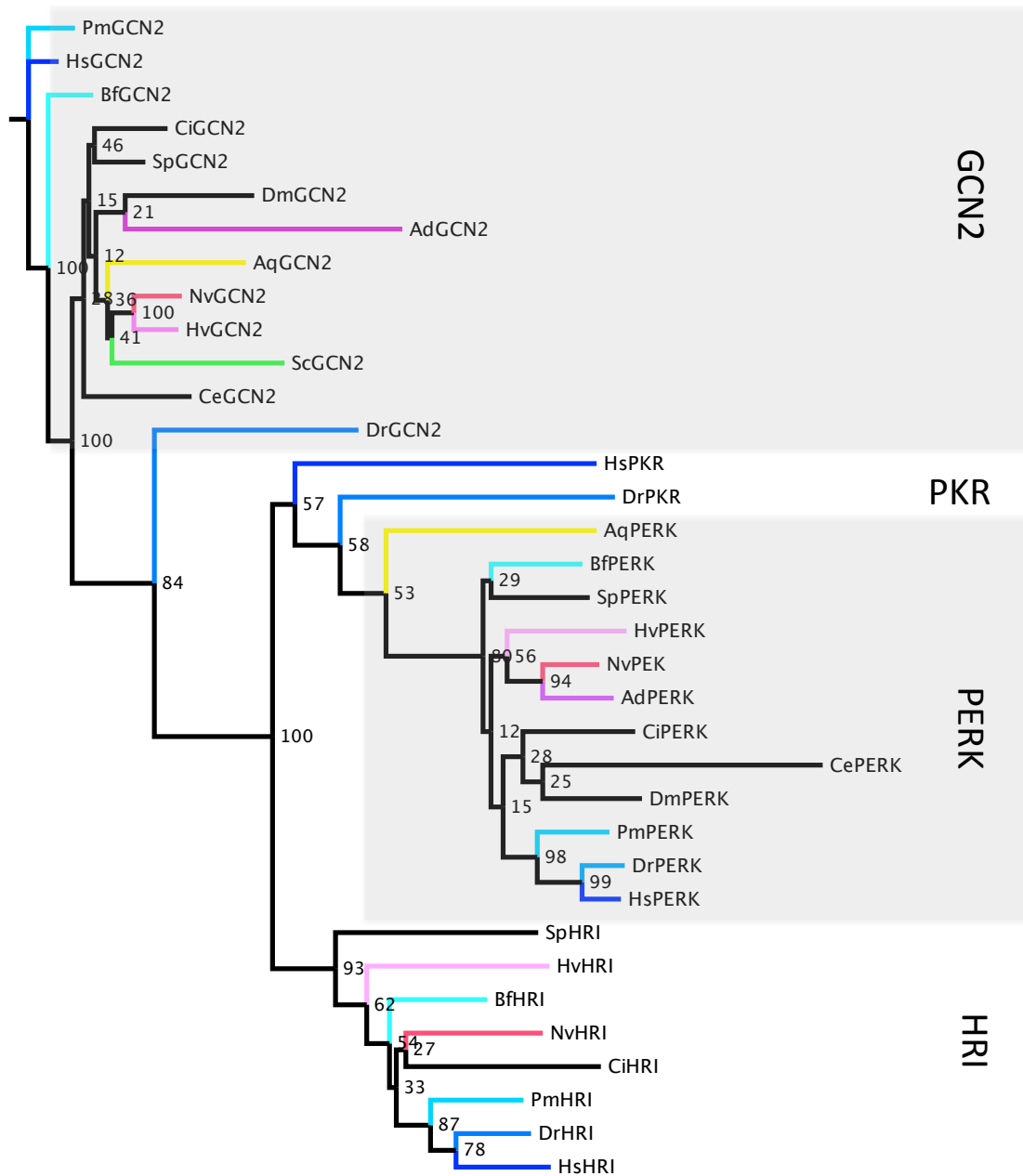
**Figure 6 | HF increases phosphorylation of eIF2α and induces ASNS and HERP gene expression in juvenile *Nematostella*.** (left panel) Juvenile (1 month old, not yet fed) *Nematostella* were treated with or without 10 uM HF for 4 hours and assayed by immunoblot for phosphorylated and total eIF2α. (right panel) mRNA expression of ASNS and HERP from juveniles under the same conditions as on the left. Densitometry data are displayed below the blot and presented as the mean ratio of phosphorylated eIF2α/total eIF2α ± s.d. from duplicate biological replicates. mRNA expression was normalized to expression of 18S mRNA and is represented as the mean of duplicate biological replicates ± s.d. Data are representative of two (left) and three (right) separate experiments. Confidence intervals (*P* value) for the effect of HF versus untreated (DMSO) were determined using a one tailed Student's t-test. All blots were also probed for actin to confirm equal loading and that total eIF2α levels did not change with treatment. (not shown).



**Figure 7 | Temperature, increased salinity, and copper exposure induce ISR readouts in adult *Nematostella*.** (a-d) Temperature. (a) Phosphorylated and total eIF2α levels assayed by immunoblot after exposure to room temperature (22°C) or 37°C for durations ranging between 30 minutes to 6 hours indicated in figure. (b) Phosphorylated and total eIF2α levels assayed by immunoblot after exposure to 37°C for durations ranging between 10 minutes and 2 hours. (c) Phosphorylated and total eIF2α levels assayed by immunoblot after exposure for 30 minutes to low 4°C and warm 37°C for 30 minutes. (d) mRNA expression of ASNS, HERP, and BiP 3.5 hours after 30 minute exposure to 37°C. (e) Salinity. mRNA expression of ASNS and HERP after exposure to 0, 13, or 50 ppt artificial seawater for 4 hours. (f) Copper. mRNA expression of ASNS and HERP after exposure to 20 μM copper for 4 hours. (a-c) Densitometry data are displayed below the blot and presented as the mean ratio of phosphorylated eIF2α/total eIF2α ± s.d. from triplicate biological replicates. mRNA expression was normalized to expression of 18S mRNA and are represented as the mean of biological triplicates ± s.d. Data in (a-c, f) are representative of one experiment and data in (d,e) are representative of two separate experiments. Confidence intervals (*P* value) for the effect of treatment versus untreated were determined using a one tailed Student's *t*-test. All blots were also probed for actin to confirm equal loading and that total eIF2α levels did not change with treatment. (not shown).



# SUPPLEMENTARY FIGURES AND TABLES

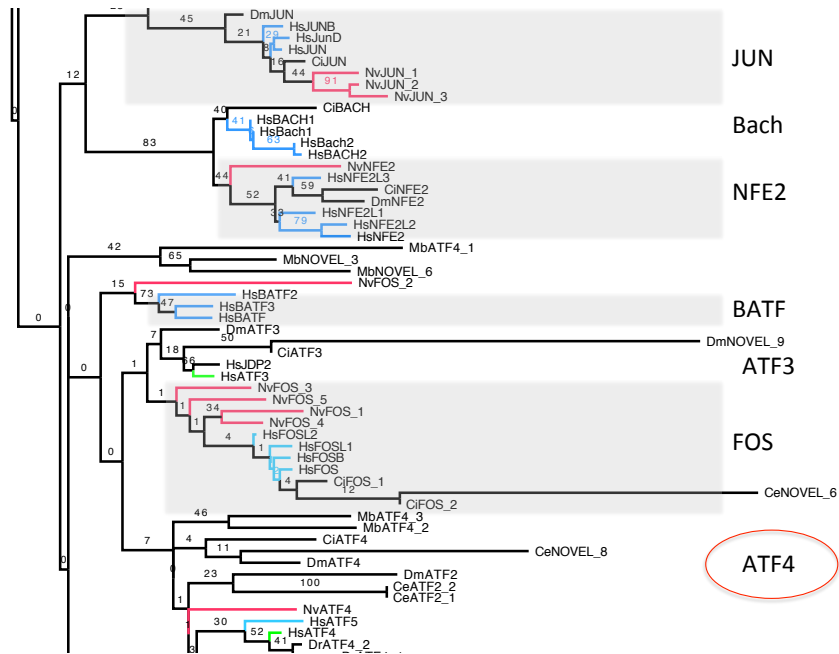


**Figure S1 | Phylogram of maximum likelihood analysis of eIF2 $\alpha$  kinases from 13 eukaryotic genomes.** *Saccharomyces cerevisiae* (Sc), *Amphimedon queenslandica* (Aq), *Acropora digitifera* (Ad), *Nematostella vectensis* (Nv), *Hydra vulgaris* (Hv), *Caenorhabditis elegans* (Ce), *Drosophila melanogaster* (Dm), *Ciona intestinalis* (Ci), *Strongylocentrotus purpuratus* (Sp), *Branchiostoma floridae* (Bf), *Petromyzon marinus* (Pm), *Danio rerio* (Dr), and *Homo sapiens* (Hs). Blue shaded lines indicate vertebrate species, black lines indicate non-vertebrate deuterostomes and protostomes, red shaded lines indicate cnidarian species, yellow indicates sponge, and green indicates yeast. Node labels represent percentage of 1,000 bootstraps and labels on the far right indicate kinase type.

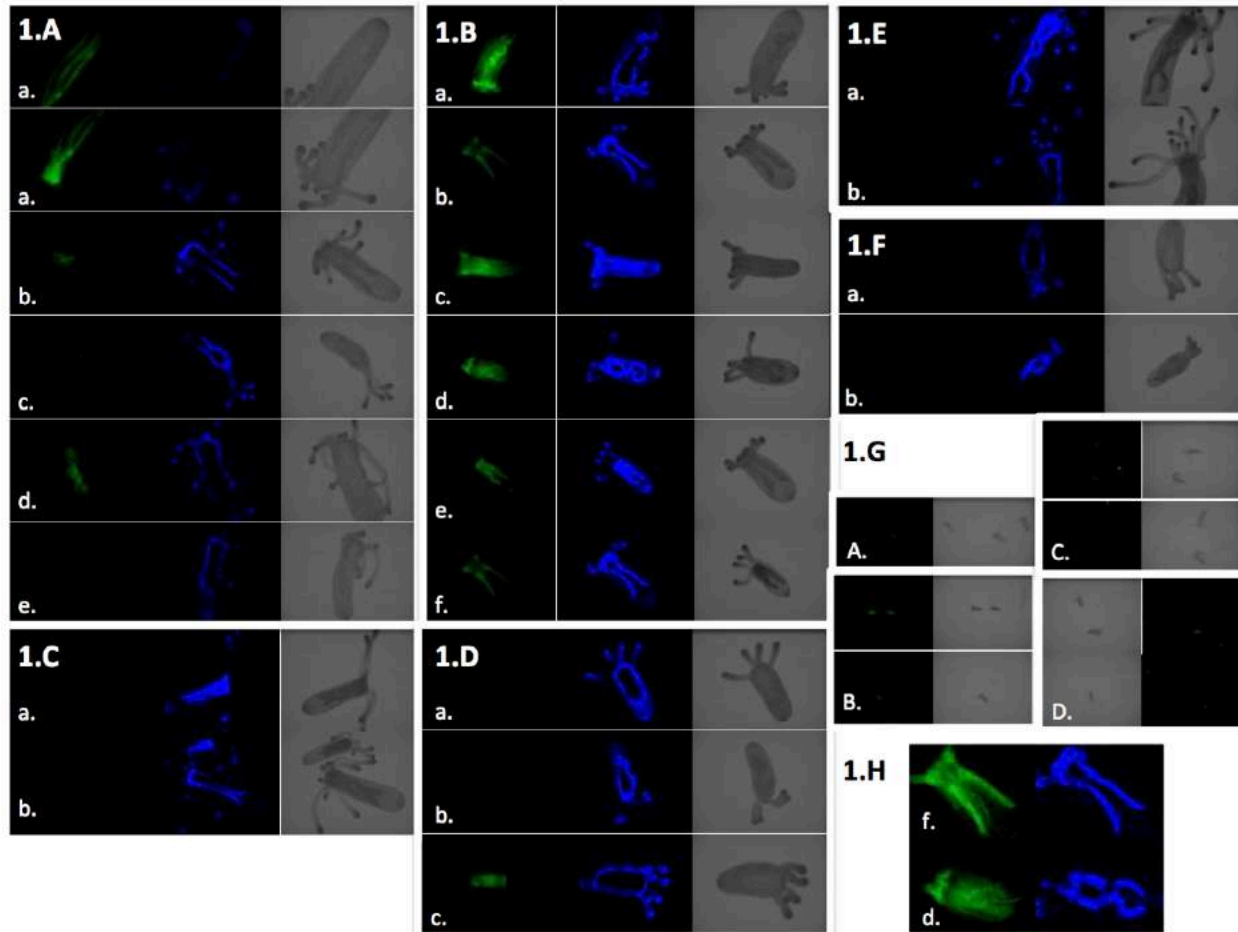
A.



B.







**Figure S3 | Localization of peIF2α in juvenile *Nematostella*.** Whole mount immunofluorescence was performed with peIF2α antibody on juveniles treated with or without 500 uM DTT for 30 minutes. Each row within each panel represents a single animal. peIF2α was probed with FITC secondary (green) and nuclei are DAPI stained (blue). (a) No DTT treatment. (b) DTT treatment. (c) No DTT treatment or primary antibody. (d) DTT treatment and no primary antibody. (e) No DTT treatment or secondary antibody. (f) DTT treatment and no secondary. (g) Low magnification. (h) High magnification of individuals from panel (b) showing cytoplasmic staining.

ISR Protein	Human NCBI Accession #:	<i>Nematostella</i> JGI ID:	Tulin et. al Transcriptome Identifier:	<i>Nematostella</i> sequence used in study	Additional Species Mentioned:
elf2α	NP_004085.1	185797		JGI	
GCN1	NP_006827.1	168356, 20230	comp4420_c0_seq1, comp8202_c0_seq1	concatenated Tulin	<i>S. cerevisiae</i> : YDR283V, <i>D. melanogaster</i> : FBpp0085148, <i>H. vulgaris</i> : XP_002154172.2, <i>A. digitifera</i> : aug_v2a.04771.t1, <i>C. elegans</i> : DC2.7a, <i>C. intestinalis</i> : ENSCINP00000035877, <i>S. purpuratus</i> : XP_785160.3, <i>B. floridae</i> : XP_002613455.1,
GCN2	EAW92388.1	30010		JGI	<i>P. marinus</i> : ENSPMAP00000007704, <i>D. rerio</i> : ENSDARP00000098667
PERK	AAF61199.1	238489		JGI	<i>A. queenslandica</i> : PAC:15713749, <i>H. vulgaris</i> : XP_004208197.1, <i>A. digitifera</i> : aug_v2a.06243.t1, <i>C. elegans</i> : F46C3.1, <i>C. intestinalis</i> : ENSCINP00000006293, <i>S. purpuratus</i> : XP_792135.3, <i>B. floridae</i> : XP_002603649.1, <i>P. marinus</i> : ENSPMAP00000006894, <i>D. rerio</i> : ENSDARP00000120826, <i>D. melanogaster</i> : FBpp0078417
HRI	AAF70289.1	200396		JGI	<i>H. vulgaris</i> : XP_002158812.1, <i>S. purpuratus</i> : XP_003728812.1, <i>B. floridae</i> : XP_002597671.1, <i>C. intestinalis</i> : ENSCINP00000025661, <i>D. rerio</i> : ENSDARP00000002313, <i>P. marinus</i> : ENSPMAP00000002333
PKR	P19525.2	-	-	-	<i>D. rerio</i> : ENSDARP00000090291
ATF4	NP_001666.2	sca303	comp3036_c0_seq1	Tulin	<i>T. adhaerens</i> : TriadT63596, <i>A. queenslandica</i> : PAC:15724215
CHOP	P35638.1	-	-	-	-
GADD34	O75807.1	238227		JGI	
CreP	NP_116222.4	-	-	-	-
SNAT2	Q96QD8.2	83628	comp4370_c0_seq2	Tulin	
xCT	Q9UPY5.1	167234		JGI	
TRB3	Q96RU7.2	164837		JGI	
CHAC1	Q98UX1.2	208105		JGI	
BiP (GRP78)	P11021.2	216823		JGI	<i>D. melanogaster</i> : NP_611863.1
ASNS	AAA52756.1	163815		JGI	
HERP1	Q15011.1	29773	comp8577_c0_seq1	Tulin	
CAT-1	P30825.1	-	-	-	-
VEGF	AAA35789.1	-	-	-	-
REDD1	Q9NX09.1	-	-	-	-

**Table S1 | Sequences used in study.**

Gene	Primer	Primer Sequence
NvASNS	forward	TGTAGCTGCTGCACGTAAGG
	reverse	GCAACCAAGGTACATCGGAAC
NvHERP	forward	CATAACATGGGATTGCCACA
	reverse	GTTGCTGTTGACAGGAGGT
NvCHAC1	forward	TTGCAGATAGTGTCTGGCTCA
	reverse	TTCGACTTGAGGGACTTTGG
NvTRIB3	forward	TGTCACAGTTCGGGTGTTGT
	reverse	GCAGCCATGTTTGTCTCTCA
NvSNAT2	forward	CAGCATCAACCTCAGCCTTC
	reverse	GAGCAATCGTCCAAACCACT
NvxCT	forward	AAACTTCTTGCGGCTGTCTG
	reverse	AAGCTGCCAGTGAAACCTTG
NvBiP	forward	GAGCTCTCAGTACACAGCACCA
	reverse	ACTTTCTGGACGGGCTTCA
NvActin	forward	TACGAGTTGCCTGATGGACA
	reverse	TACCCAGAAAGGAGGGCTGA
Nv18S*	forward	GACTCAACACGGGGAAACTC
	reverse	GCACCACCATAGTAATC

\* sequences from Tarrant Lab

**Table S2 | Primer sequences used in study.**

## REFERENCES

1. Spriggs KA, Bushell M, & Willis AE (2010) Translational regulation of gene expression during conditions of cell stress. *Molecular cell* 40(2):228-237.
2. Holcik M & Sonenberg N (2005) Translational control in stress and apoptosis. *Nature reviews. Molecular cell biology* 6(4):318-327.
3. Yamasaki S & Anderson P (2008) Reprogramming mRNA translation during stress. *Current opinion in cell biology* 20(2):222-226.
4. Kilberg MS, Shan J, & Su N (2009) ATF4-dependent transcription mediates signaling of amino acid limitation. *Trends in endocrinology and metabolism: TEM* 20(9):436-443.
5. B'Chir W, *et al.* (2013) The eIF2alpha/ATF4 pathway is essential for stress-induced autophagy gene expression. *Nucleic Acids Res* 41(16):7683-7699.
6. Jiang HY, *et al.* (2003) Phosphorylation of the Subunit of Eukaryotic Initiation Factor 2 Is Required for Activation of NF- B in Response to Diverse Cellular Stresses. *Molecular and cellular biology* 23(16):5651-5663.
7. Harding HP, *et al.* (2003) An Integrated Stress Response Regulates Amino Acid Metabolism and Resistance to Oxidative Stress. *Molecular cell* 11:619-633.
8. Sonenberg N & Hinnebusch AG (2009) Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. *Cell* 136:731-745.
9. Donnelly N, Gorman AM, Gupta S, & Samali A (2013) The eIF2alpha kinases: their structures and functions. *Cellular and molecular life sciences : CMLS* 70(19):3493-3511.
10. Castilho BA, *et al.* (2014) Keeping the eIF2 alpha kinase Gcn2 in check. *Biochimica et biophysica acta*.
11. Hernandez G, Altmann M, & Lasko P (2010) Origins and evolution of the mechanisms regulating translation initiation in eukaryotes. *Trends in biochemical sciences* 35(2):63-73.
12. Lageix S, *et al.* (2008) Arabidopsis eIF2alpha kinase GCN2 is essential for growth in stress conditions and is activated by wounding. *BMC plant biology* 8:134.
13. Tarrant AM, Reitzel AM, Kwok CK, & Jenny MJ (2014) Activation of the cnidarian oxidative stress response by ultraviolet light, polycyclic aromatic hydrocarbons and crude oil. *The Journal of experimental biology* 217:1444-1453.
14. Reitzel AM, *et al.* (2013) Physiological and developmental responses to temperature by the sea anemone *Nematostella vectensis*. *Marine Ecology Progress Series* 484:115-130.
15. Darling JA, *et al.* (2005) Rising starlet: the starlet sea anemone, *Nematostella vectensis*. *Bioessays* 27(2):211-221.
16. Technau U & Steele RE (2011) Evolutionary crossroads in developmental biology: Cnidaria. *Development* 138:1447-1458.
17. Rentzsch F, Fritzenwanker JH, Scholz CB, & Technau U (2008) FGF signalling controls formation of the apical sensory organ in the cnidarian *Nematostella vectensis*. *Development* 135(10):1761-1769.
18. Genikhovich G & Technau U (2009) The starlet sea anemone *Nematostella vectensis*: an anthozoan model organism for studies in comparative genomics and functional evolutionary developmental biology. *Cold Spring Harbor protocols* 2009(9):pdb emo129.
19. Putnam NH, *et al.* (2007) Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science* 317(5834):86-94.
20. Ikmi A & Gibson MC (2010) Identification and in vivo characterization of NvFP-7R, a developmentally regulated red fluorescent protein of *Nematostella vectensis*. *PloS one* 5(7):e11807.
21. Rutkowski DT & Kaufman RJ (All roads lead to ATF4. *Developmental Cell*:442-444.
22. Baird TD & Wek RC (2012) Eukaryotic initiation factor 2 phosphorylation and translational control in metabolism. *Advances in nutrition* 3(3):307-321.

23. Kroemer G, Marino G, & Levine B (2010) Autophagy and the Integrated Stress Response. *Molecular cell* 40:280-293.
24. Wek RC, Jiang H-Y, & Anthony TG (2006) Coping with stress: eIF2 kinases and translational control. *Biochemical Society Transactions* 34:7-11.
25. Wek RC & Cavener DR (2007) Translational control and the unfolded protein response. *Antioxidants & redox signaling* 9(12):2357-2371.
26. Barbosa C, Peixeiro I, & Romao L (2013) Gene expression regulation by upstream open reading frames and human disease. *PLoS genetics* 9(8):e1003529.
27. Lu PD, Harding HP, & Ron D (2004) Translation reinitiation at alternative open reading frames regulates gene expression in an integrated stress response. *Journal of Cell Biology* 167(1):27-33.
28. Vattam KM & Wek RC (2004) Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. *Proceedings of the National Academy of Sciences* 101(31):11269-11274.
29. Hinnebusch AG (2005) Translational regulation of *GCN4* and the general amino acid control of yeast. *Annual Review of Microbiology* 59:407-450.
30. Patil CK, Li H, & Walter P (2004) Gcn4p and novel upstream activating sequences regulate targets of the unfolded protein response. *PLoS biology* 2(8):E246.
31. Acharya P, Chen JJ, & Correia MA (2010) Hepatic heme-regulated inhibitor (HRI) eukaryotic initiation factor 2alpha kinase: a protagonist of heme-mediated translational control of CYP2B enzymes and a modulator of basal endoplasmic reticulum stress tone. *Molecular pharmacology* 77(4):575-592.
32. Andreev DE, *et al.* (2015) Translation of 5' leaders is pervasive in genes resistant to eIF2 repression. *eLife* 4:1-21.
33. Neafsey DE & Galagan JE (2007) Dual modes of natural selection on upstream open reading frames. *Molecular biology and evolution* 24(8):1744-1751.
34. Taniuchi S, Miyake M, Tsugawa K, Oyadomari M, & Oyadomari S (2016) Integrated stress response of vertebrates is regulated by four eIF2alpha kinases. *Scientific reports* 6:32886.
35. Natarajan K, *et al.* (2001) Transcriptional profiling shows that Gcn4p is a master regulator of gene expression during amino acid starvation in yeast. *Molecular and cellular biology* 21(13):4347-4368.
36. Harding HP, Zhang Y, & Ron D (1999) Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase. *Nature* 397:271-274.
37. Zhan K, *et al.* (2002) Phosphorylation of Eukaryotic Initiation Factor 2 by Heme-Regulated Inhibitor Kinase-Related Protein Kinases in *Schizosaccharomyces pombe* Is Important for Resistance to Environmental Stresses. *Molecular and cellular biology* 22(20):7134-7146.
38. Zhan K, Narasimhan J, & Wek RC (2004) Differential activation of eIF2 kinases in response to cellular stresses in *Schizosaccharomyces pombe*. *Genetics* 168(4):1867-1875.
39. Chen J-J & London IM (1995) Regulation of protein synthesis by heme-regulated eIF-2alpha kinase. *TIBS* 20:105-108.
40. Matts RL, Schatz JR, Hurst R, & Kagen R (1991) Toxic heavy metal ions activate the heme-regulated eukaryotic initiation factor-2alpha kinase by inhibiting the capacity of hemin-supplemented reticulocyte lysates to reduce disulfide bonds. *Journal of Biological Chemistry* 266:12695-12702.
41. Hinnebusch AG (1994) Translational control of GCN4: an *in vivo* barometer of initiation-factor activity. *TIBS* 19:410-414.
42. Silverman RH & Williams BRG (1999) Translational control perks up. *Nature* 397:208-211.
43. Ameri K & Harris AL (2008) Activating transcription factor 4. *The international journal of biochemistry & cell biology* 40(1):14-21.
44. Siu F, Bain PJ, LeBlanc-Chaffin R, Chen H, & Kilberg MS (2002) ATF4 is a mediator of the nutrient-sensing response pathway that activates the human asparagine synthetase gene. *The Journal of biological chemistry* 277(27):24120-24127.

45. Balasubramanian MH, Butterworth EA, & Kilberg MS (2013) Asparagine synthetase: regulation by cell stress and involvement in tumor biology. *Am J Endocrinol Metab* 304:789-799.
46. Barbosa-Tessmann IP, *et al.* (1999) Activation of the Unfolded Protein Response Pathway Induces Human Asparagine Synthetase Gene Expression. *The Journal of biological chemistry* 274(44):31139-31144.
47. Gaccioli F, *et al.* (2006) Amino acid starvation induces the SNAT2 neutral amino acid transporter by a mechanism that involves eukaryotic initiation factor 2alpha phosphorylation and cap-independent translation. *The Journal of biological chemistry* 281(26):17929-17940.
48. Hyde R, Cwiklinski EL, MacAulay K, Taylor PM, & Hundal HS (2007) Distinct sensor pathways in the hierarchical control of SNAT2, a putative amino acid transceptor, by amino acid availability. *The Journal of biological chemistry* 282(27):19788-19798.
49. Lewerenz J & Maher P (2009) Basal levels of eIF2alpha phosphorylation determine cellular antioxidant status by regulating ATF4 and xCT expression. *The Journal of biological chemistry* 284(2):1106-1115.
50. Lewerenz J, *et al.* (2013) The cystine/glutamate antiporter system x(c)(-) in health and disease: from molecular mechanisms to novel therapeutic opportunities. *Antioxidants & redox signaling* 18(5):522-555.
51. Hori O, *et al.* (2004) Role of Herp in the endoplasmic reticulum stress response. *Genes to cells : devoted to molecular & cellular mechanisms* 9(5):457-469.
52. Huang CH, Chu YR, Ye Y, & Chen X (2014) Role of HERP and a HERP-related protein in HRD1-dependent protein degradation at the endoplasmic reticulum. *The Journal of biological chemistry* 289(7):4444-4454.
53. Kokame K, Agarwala KL, Kato H, & Miyata T (2000) Herp, a new ubiquitin-like membrane protein induced by endoplasmic reticulum stress. *The Journal of biological chemistry* 275(42):32846-32853.
54. Quiroga C, *et al.* (2013) Herp depletion protects from protein aggregation by up-regulating autophagy. *Biochimica et biophysica acta* 1833(12):3295-3305.
55. Crawford RR, *et al.* (2015) Human CHAC1 Protein Degrades Glutathione and mRNA Induction is Regulated by the Transcription Factors ATF4 and ATF3 and a Bipartite ATF/CRE Element. *The Journal of biological chemistry* 290(25):15878-15891.
56. Novoa I, Zeng H, Harding HP, & Ron D (2001) Feedback Inhibition of the Unfolded Protein Response by GADD34-mediated Dephosphorylation of eIF2alpha. *The Journal of cell biology* 153:1011-1021.
57. Ord D & Ord T (2005) Characterization of human NIPK (TRB3, SKIP3) gene activation in stressful conditions. *Biochemical and biophysical research communications* 330(1):210-218.
58. Ohoka N, Yoshii S, Hattori T, Onozaki K, & Hayashi H (2005) TRB3, a novel ER stress-inducible gene, is induced via ATF4-CHOP pathway and is involved in cell death. *The EMBO journal* 24:1243-1255.
59. Hegedus Z, Czibula A, & Kiss-Toth E (2006) Tribbles: novel regulators of cell function; evolutionary aspects. *Cellular and molecular life sciences : CMLS* 63(14):1632-1641.
60. Dang V-D, Valens M, Bolotin-Fukuhara M, & Daignan-Fornier B (1996) Cloning of the ASN1 and ASN2 genes encoding asparagine synthetases in *Saccharomyces cerevisiae*: differential regulation by the CCAAT-box-binding factor. *Molecular Microbiology* 22:681-692.
61. Ramos F & Wiame J-M (1980) Two Asparagine Synthetases in *Saccharomyces cerevisiae*. *European Journal of Biochemistry* 108:373-377.
62. Kennish MJ (2002) Environmental threats and environmental future of estuaries. *Environmental Conservation*:78-107.
63. Reitzel AM, Sullivan JC, Traylor-Knowles N, & Finnerty JR (2008) Genomic Survey of Candidate Stress-Response Genes in the Estuarine Anemone *Nematostella vectensis*. *Biological Bulletin* 214:233-254.



64. Doney SC, *et al.* (2012) Climate change impacts on marine ecosystems. *Annual review of marine science* 4:11-37.
65. Abele D, Burlando B, Viarengo A, & Portner H-O (1998) Exposure to elevated temperatures and hydrogen peroxide elicits oxidative stress and antioxidant response in the Antarctic intertidal limpet *Nacella concinna*. *Comparative Biochemistry and Physiology* 120:425-435.
66. Abele-Oeschger D, Sartoris FJ, & Portner H-O (1997) Hydrogen Peroxide Causes a Decrease in Aerobic Metabolic Rate and in Intracellular pH in the Shrimp *Crangon crangon*. *Comparative Biochemistry and Physiology* 117(123-129).
67. Lee EC-H & Strange K (2012) GCN-2 dependent inhibition of protein synthesis activates osmosensitive gene transcription via WNK and Ste20 kinase signaling. *American Journal of Physiology - Cell Physiology* 303:C1269-C1277.
68. Yang D, *et al.* (2014) Hypoxia triggers endothelial endoplasmic reticulum stress and apoptosis via induction of VLDL receptor. *FEBS letters* 588(23):4448-4456.
69. Moenner M, Pluquet O, Bouchecareilh M, & Chevet E (2007) Integrated endoplasmic reticulum stress responses in cancer. *Cancer research* 67(22):10631-10634.
70. Komoike Y & Matsuoka M (2013) Exposure to tributyltin induces endoplasmic reticulum stress and the unfolded protein response in zebrafish. *Aquatic toxicology* 142-143:221-229.
71. Bettaieb A & Averill-Bates DA (2015) Thermotolerance induced at a mild temperature of 40 degrees C alleviates heat shock-induced ER stress and apoptosis in HeLa cells. *Biochimica et biophysica acta* 1853(1):52-62.
72. Tulin S, Aguiar D, Istrail S, & Smith J (2013) A quantitative reference transcriptome for *Nematostella vectensis* early embryonic development: a pipeline for de novo assembly in emerging model systems. *EvoDevo* 4(1):16.
73. Dar AC, Dever TE, & Sicheri F (2005) Higher-order substrate recognition of eIF2alpha by the RNA-dependent protein kinase PKR. *Cell* 122(6):887-900.
74. Wek SA, Zhu S, & Wek RC (1995) The Histidyl-tRNA Synthetase-Related Sequence in the eIF-2alpha Protein Kinase GCN2 Interacts with tRNA and Is Required for Activation in Response to Starvation for Different Amino Acids. *Molecular and cellular biology* 15(8):4497-4506.
75. Pomar N, *et al.* (2003) Functional characterization of *Drosophila melanogaster* PERK eukaryotic initiation factor 2alpha (eIF2alpha) kinase. *European Journal of Biochemistry* 270(2):293-306.
76. Igarashi J, *et al.* (2008) Elucidation of the heme binding site of heme-regulated eukaryotic initiation factor 2alpha kinase and the role of the regulatory motif in heme sensing by spectroscopic and catalytic studies of mutant proteins. *The Journal of biological chemistry* 283(27):18782-18791.
77. Rafie-Kolpin M, *et al.* (2000) Two heme-binding domains of heme-regulated eukaryotic initiation factor-2alpha kinase. *Journal of Biological Chemistry* 275:5171-5178.
78. Ishikawa T, *et al.* (2013) ATF6alpha/beta-mediated adjustment of ER chaperone levels is essential for development of the notochord in medaka fish. *Molecular biology of the cell* 24(9):1387-1395.
79. Wang C & Guo F (2012) Effects of activating transcription factor 4 deficiency on carbohydrate and lipid metabolism in mammals. *IUBMB life* 64(3):226-230.
80. Reinke AW, Baek J, Ashenberg O, & Keating AE (2013) Networks of bZIP protein-protein interactions diversified over a billion years of evolution. *Science* 340(6133):730-734.
81. Jousse C, *et al.* (2003) Inhibition of a constitutive translation initiation factor 2alpha phosphatase, CREP, promotes survival of stressed cells. *The Journal of cell biology* 163(4):767-775.
82. Malzer E, *et al.* (2013) Coordinate regulation of eIF2alpha phosphorylation by PPP1R15 and GCN2 is required during *Drosophila* development. *Journal of cell science* 126(Pt 6):1406-1415.
83. DuRose JB, Tam AB, & Niwa M (2006) Intrinsic capacities of molecular sensors of the unfolded protein response to sense alternate forms of endoplasmic reticulum stress. *Molecular biology of the cell* 17(7):3095-3107.

84. Keller TL, *et al.* (2012) Halofuginone and other febrifugine derivatives inhibit prolyl-tRNA synthetase. *Nature Chemical Biology* 8:311-317.
85. Ray PS, *et al.* (2011) Evolution of function of a fused metazoan tRNA synthetase. *Molecular biology and evolution* 28(1):437-447.
86. Kim H & Strange K (2013) Changes in translation rate modulate stress-induced damage of diverse proteins. *American journal of physiology. Cell physiology* 305(12):C1257-1264.
87. Goossens A, Dever TE, Pascual-Ahuir A, & Serrano R (2001) The protein kinase Gcn2p mediates sodium toxicity in yeast. *The Journal of biological chemistry* 276(33):30753-30760.
88. Hand C & Uhlinger KR (1994) The Unique, Widely Distributed, Estuarine Sea Anemone, *Nematostella vectensis* Stephenson: A Review, New Facts, and Questions. *Estuaries* 17(2):501-508.
89. Lee HC, *et al.* (2014) Zebrafish transgenic line huORFZ is an effective living bioindicator for detecting environmental toxicants. *PloS one* 9(3):e90160.
90. Gotze S, Matoo OB, Beniash E, Saborowski R, & Sokolova IM (2014) Interactive effects of CO<sub>2</sub> and trace metals on the proteasome activity and cellular stress response of marine bivalves *Crassostrea virginica* and *Mercenaria mercenaria*. *Aquatic toxicology* 149:65-82.
91. Tardito S, *et al.* (2011) Copper binding agents acting as copper ionophores lead to caspase inhibition and paraptotic cell death in human cancer cells. *Journal of the American Chemical Society* 133:6235-6242.
92. Quiroz N, *et al.* (2015) Transcriptional activation of glutathione pathways and role of glucose homeostasis during copper imbalance. *Biometals* 28:321-328.
93. Sonnack L, *et al.* (2015) Effects of metal exposure on motor neuron development, neuromasts and the escape response of zebrafish embryos. *Neurotoxicology and teratology* 50:33-42.
94. Lovejoy DB & Guillemin GJ (2014) The potential for transition metal-mediated neurodegeneration in amyotrophic lateral sclerosis. *Frontiers in aging neuroscience* 6:173.
95. Chaveroux C, *et al.* (2015) In vivo imaging of the spatiotemporal activity of the eIF2 $\alpha$ -ATF4 signaling pathway: Insights into stress and related disorders. *Science Signaling* 8(374).
96. Kang K, Ryoo HD, Park JE, Yoon JH, & Kang MJ (2015) A Drosophila Reporter for the Translational Activation of ATF4 Marks Stressed Cells during Development. *PloS one* 10(5):e0126795.
97. Frerichs KU, *et al.* (1998) Suppression of protein synthesis in brain during hibernation involves inhibition of protein initiation and elongation. *Proceedings of the National Academy of Sciences* 95:14511-14516.
98. Flicek P, *et al.* (2014) Ensembl 2014. *Nucleic Acids Res* 42(Database issue):D749-755.
99. Shinzato C, *et al.* (2011) Using the *Acropora digitifera* genome to understand coral responses to environmental change. *Nature* 476:320-324.
100. Ltd B (2012) Geneious Pro 5.6.6.
101. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32(5):1792-1797.
102. Abascal F, Zardoya R, & Posada D (2009) ProtTest: Selection of best-fit models of protein evolution version 2.4. *Bioinformatics* 21:2104-2105.
103. Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22(21):2688-2690.
104. Rambaut A (2012) Tree Figure Drawing Tool Version 1.4.0. *Institute of Evolutionary Biology, University of Edinburgh*.
105. Fritzenwanker JH & Technau U (2002) Induction of gametogenesis in the basal cnidarian *Nematostella vectensis* (Anthozoa). *Development genes and evolution* 212(2):99-103.
106. Stefanik DJ, Friedman LE, & Finnerty JR (2013) Collecting, rearing, spawning and inducing regeneration of the starlet sea anemone, *Nematostella vectensis*. *Nature protocols* 8(5):916-923.

107. Wolenski FS, Layden MJ, Martindale MQ, Gilmore TD, & Finnerty JR (2013) Characterizing the spatiotemporal expression of RNAs and proteins in the starlet sea anemone, *Nematostella vectensis*. *Nature protocols* 8(5):900-915.



## **CHAPTER 3**

**The evolutionarily conserved protein GCN1 is a novel branch point in the mammalian AAR pathway**



## INTRODUCTION

### **Nutrient sensing pathways have evolved roles in cellular protection**

The ability to successfully respond to environmental stimuli, such as changing nutrient conditions, is essential to an organism's survival. Signaling pathways that sense nutrient availability are also integral to cellular defense and tissue homeostasis (1-4). A well-recognized example of adaptation to nutrient limitation occurs in bacteria and utilizes the protein synthetic machinery for signal generation. Stringent control, first described in *E. coli*, but also found in plant plastids, detects amino acid depletion by sensing non-aminoacylated, or uncharged, tRNA and responds by comprehensive transcriptional reprogramming to induce growth arrest, upregulate survival genes and increase virulence in pathogenic bacteria (5, 6). Induction of nutrient sensing pathways is a phylogenetically conserved feature in the control of lifespan in whole organisms such as yeast, *C. elegans*, and *Drosophila*, and plays a key role in the control of aging and age-related diseases (7, 8). Amino acid sensing pathways, like mTOR and the AAR pathway, also participate in immuno-regulation (9, 10). The mammalian immune system protects tissue and organ structural integrity by balancing protective inflammatory response to a stress with its timely resolution. During injury or infection, tissue and immune homeostasis are shifted; failure to restore balance can result in sustained inflammatory cytokine signaling, pathologic inflammation and tissue damage (11). Mammalian immune regulation has evolved to utilize the induced breakdown of amino acids through transcriptional upregulation of amino acid-degrading enzymes (and the concomitant induction of the AAR pathway), to limit hyper-inflammatory innate responses and promote tolerance to self (2, 12, 13).

### **HF's therapeutic effects depend upon activation of the AAR pathway**

Our lab became interested in the connection between the activation of nutrient stress response pathways and the control of tissue inflammation when the lab solved the molecular target of the small molecule therapeutic, HF (14). HF inhibits the catalytic activity of glutamyl-prolyl-tRNA synthetase (EPRS), thereby activating the nutrient-sensing AAR pathway (14, 15). Under conditions of amino acid limitation, aaRSs, the cellular enzymes that charge amino acids into cognate tRNA, cannot perform their function and non-aminoacylated or uncharged tRNA accumulates. By inhibiting the PRS activity of EPRS, HF acts as amino acid restriction mimetic, using the cell's protein synthetic machinery to generate the accumulation of uncharged tRNA (14). We now know that the addition of HF to cytokine-stimulated cells selectively inhibits inflammatory responses in diverse cell types, and that these cellular responses to HF treatment can occur in cells that lack the signature AAR pathway effector GCN2 (Kim et al, manuscript in preparation). Additionally, we know that borrelidin, a TRS inhibitor, can reproduce many of the key therapeutic tissue responses of HF treatment (Kim et al, manuscript in preparation). The ability of these inhibitors to induce therapeutic effects in cells that lack GCN2 indicates both the existence of a previously unrecognized branch point in the mammalian AAR pathway, and the presence of a novel sensor of uncharged tRNA in mammalian cells.

### **HF inhibits an inflammatory program involved in chronic inflammatory diseases and RA**

HF inhibits an inflammatory program involved in chronic inflammatory disease and rheumatoid arthritis (RA), including inhibition of pro-inflammatory T-cell differentiation and transcription of pro-inflammatory mediators (16) (Kim et al, manuscript in preparation). An essential component of any protective immune response is proper regulation, which includes downregulation of the immune response. In chronic inflammatory disease, like RA, immune responses are dysregulated and a positive feedback loop of signals activates and perpetuates a tissue-damaging program (11). T helper 17 (Th17) cells, a pro-inflammatory T-cell subset, are broad mediators of autoimmunity and implicated in RA (17, 18). Via AAR induction, HF inhibits Th17 differentiation, which can be rescued by addition of proline and mimicked by amino acid depletion (16). RA is a chronic autoimmune inflammatory disease that primarily attacks synovial joints exhibiting synovial inflammation, proliferation of synovial tissues, and destruction of cartilage and bone. Fibroblast-like synoviocytes (FLS), resident mesenchymal cells of the synovium (the tissue lining of the edges of the joint space), play an important role in the pathogenesis of

RA (19, 20). In RA, activation of FLS by cytokines and other soluble factors leads to FLS production of molecules that recruit, retain, and activate cells of the immune system and other joint cells, leading to the promotion of ongoing inflammation and tissue damage (20). Activated FLS further contribute to tissue damage by the upregulation and secretion of matrix degrading enzymes. Our lab has found that HF treatment significantly downregulates transcription of many of these pro-inflammatory and tissue remodeling mediators in both TNF- $\alpha$  activated FLS and other cytokine-stimulated cell types implicated in inflammatory and fibrotic disease (Kim et al, manuscript in preparation).

These observations point to a GCN2-independent pathway induced by uncharged tRNA in mammalian cells but what proteins mediate this pathway are unknown. The conservation of the AAR pathway across eukaryotes provides two proteins as candidate mediators of this novel pathway, GCN1 and ABCF3. In yeast these proteins function upstream of GCN2: where GCN1 is required for GCN2 activation by uncharged tRNA and the yeast ortholog of ABCF3, GCN20, is required for full AAR pathway induction (21-27). In the nematode, *C. elegans*, GCN1 and ABCF3 mediate apoptotic functions independently of GCN2 (28). The functions of these proteins remain unknown in mammalian cells. To investigate whether GCN1 and ABCF3 represent a branch point within the AAR pathway, mediating GCN2-dependent and independent pathways I first tested whether GCN1 and ABCF3 are required for AAR signaling in mammalian cells. I did this by assessing AAR pathway readouts in HF-treated GCN1 and ABCF3 knockout cell lines. Since GCN1 but not ABCF3 was required for AAR pathway activation, I next tested whether GCN1 was required for GCN2-independent therapeutic transcriptional effects in HF-treated GCN1-knockdown FLS. This showed that GCN1 was required for GCN2-independent effects, revealing GCN1 as a branch point in the mammalian AAR pathway.

How GCN1 functions and what additional functions this GCN2-independent pathway may mediate are unknown. GCN1 is presumed to function at the ribosome due to its functional role described in yeast, but its localization has not been characterized (23). GCN1 has not been studied in mammalian cells but has been identified as a putative interactor with a large set of proteins seemingly distinct from its role at the ribosome (29-31). These interactors include proteins known to localize to the nucleus such as components of the mediator complex (31-40). These observations suggest that GCN1 may function away from the ribosome, potentially mediating GCN2-independent functions by interacting with proteins in other cell compartments, such as the nucleus. HF treatment could regulate GCN1 location, moving GCN1 from the ribosome into the nucleus. To determine whether GCN1 may be present elsewhere in the cell and if HF regulates GCN1 location, I biochemically fractionated cells treated with and without HF. Since GCN1 was present in multiple cell fractions and its distribution changed with HF, I then tried to see if GCN1 was located in the nucleus by microscopy with and without HF treatment. Additionally, I immunoprecipitated GCN1 from isolated nuclei, and identified co-precipitating proteins by mass spectrometry analysis, both to confirm previously identified GCN1 nuclear interactors and identify GCN1 nuclear interactors in an unbiased manner.

## RESULTS

### **GCN1, but not ABCF3, is required for GCN2 phosphorylation and the AAR transcriptional response in mammalian cells**

To ask whether GCN1 and ABCF3 represent a branch point in the mammalian AAR, we first needed to determine if these proteins are required for the AAR in mammals. In yeast, GCN1 and GCN20 function upstream of GCN2, and are required for GCN2 phosphorylation by uncharged tRNA (21, 22). These genes are conserved in the mammalian genome and in *C. elegans*, GCN1 and ABCF3 mutants exhibit decreased p-eIF2 $\alpha$  levels (28). Overexpression of a fragment of mammalian GCN1 analogous to the portion of yeast GCN1 that binds GCN2 in yeast, inhibits AAR activation in a mammalian cell (41). Together these observations strongly suggest that GCN1 and ABCF3 are required for the AAR in mammals, but this has not been tested. To determine if GCN1 and ABCF3 are required, CRISPR-



mediated knockout human colon carcinoma HCT116 cell lines of GCN2, GCN1, and ABCF3 were assayed for AAR induction upon HF treatment by measuring phosphorylated GCN2 and eIF2 $\alpha$  levels. In HF treated knockout GCN2 and GCN1 cell lines, GCN2 and eIF2 $\alpha$  phosphorylation was absent establishing that GCN1 is indeed required for AAR activation in mammals (**Fig 1a**). Unexpectedly, GCN2 and eIF2 $\alpha$  phosphorylation were not absent or decreased in ABCF3 knockout cells, demonstrating that ABCF3 is not required for AAR induction in mammals (**Fig 1a**). In yeast, GCN1 deletion decreases GCN20 levels, but this is not observed in GCN1 knockout HCT116 cell lines (22) (**Fig 1a**). To determine if GCN1 is required in other cell types, we knocked down GCN1 protein levels by shRNA in HeLa, the cell line used to identify putative GCN1 protein interactors in previous studies, and measured phosphorylated GCN2 and eIF2 $\alpha$  levels upon HF treatment. GCN1 knockdown HeLa also showed decreased AAR activation compared to wildtype upon HF treatment, which was rescued by ectopic expression of shRNA-resistance GCN1 (**Fig S2a**). Similarly, shRNA-mediated GCN1 knockdown in the transformed FLS cell line, K4, the cell line in which our lab has found HF regulated GCN2-independent transcriptional effects, decreased HF phosphorylation of GCN2 and eIF2 $\alpha$  (data not shown). Together these results demonstrate that GCN1 is required for the GCN2-eIF2 $\alpha$  portion of the AAR in mammals.

The GCN2-eIF2 $\alpha$ -ATF4 pathway of the AAR mediates adaptive gene transcription including amino acid biosynthetic genes and genes involved in negative feedback regulation of the pathway. To confirm whether GCN1 and ABCF3 are required for the transcriptional program of the AAR in mammals, knockout GCN2, GCN1, and ABCF3 HCT116 cell lines were assayed for mRNA expression of characteristic AAR transcriptional targets upon HF treatment. As expected, induction of TRB3, ASNS, and ATF3 by HF was absent in GCN1 null cells mirroring pGCN2 results, while in ABCF3 null cells induction by HF was unchanged relative to wildtype (**Fig 1b**), establishing that GCN1, but not ABCF3 is required for the mammalian AAR.

#### **Several HF therapeutic transcriptional effects require GCN1, but not GCN2**

GCN2 is the only known sensor of uncharged tRNA in the eukaryotic cell and the GCN2-eIF2 $\alpha$ -ATF4 pathway is the only known mediator of the uncharged tRNA signal. Our lab has found, however, that a subset of HF therapeutic effects is GCN2-independent, including transcriptional inhibition of a set pro-inflammatory mediators induced by TNF $\alpha$  in K4 FLS. Since we now know GCN1 is an upstream mediator of the AAR in mammalian cells and represents a logical candidate mediator of GCN2-independent effects, we can ask if GCN1 is required for these GCN2-independent effects. This question was first tested by Yeon-Jin Kim by testing if known GCN2-independent transcriptional effects upon HF treatment and TNF $\alpha$  stimulation were present in shRNA-mediated GCN1 knockdown K4 cells. I repeated these observations with Halofuginol (HFol), a less cytotoxic analog of HF with potential for drug development (42). HFol exhibits the same therapeutic effects of HF at similar doses, but exhibits fewer of the side effects that have hampered clinical use of HF (42) (**Fig 2b, c**). HFol treatment inhibited mRNA expression of several pro-inflammatory mediators, including CXCL10, MMP13, and IDO1, in shRNA-mediated scrambled control and GCN2 knockdown K4 cells, but not in GCN1 knockdown cells (**Fig 2a, c**). Knockdown of TRIB3 expression in GCN2 and GCN1 knockdown cells confirmed that the GCN2-eIF2 $\alpha$ -ATF4 transcriptional pathway was well knocked down in these cells (**Fig 2a**). HF inhibition of these pro-inflammatory mediators is absent with proline addition and inhibition of these pro-inflammatory mediators is recapitulated by borrelidin treatment, demonstrating that these GCN2-independent, GCN1-dependent transcriptional effects are a result of HF production of uncharged tRNA (Kim et al, manuscript in preparation). Notably, similar experiments in shRNA-mediated ABCF3 knockdown K4 cells showed that ABCF3 is not required for these effects as expected from the previous results in knockout HCT116 cells (Kim et al, manuscript in preparation). Thus GCN1 represents an upstream branch point in the mammalian AAR pathway.

#### **Several extraction methods show presence of GCN1 in cell fractions with nuclear components**

In yeast, GCN1 is known to physically associate with polysomes, and this is where GCN1's only identified function occurs with respect to GCN2, but in yeast and mammals GCN1 is putatively associated with numerous nuclear proteins (23, 25, 26). This suggests GCN1 may function elsewhere in the cell. To determine if GCN1 is present in additional cell compartments, both HeLa and K4 cells were fractionated by several methods (**Fig 3**). Fractionation using the Thermo Subcellular Protein Fractionation Kit for Cultured Cells showed that a large fraction of GCN1 is present in fractions containing nuclear components (**Fig 3a**). Digitonin extraction, which selectively permeabilizes the cell membrane extracting primarily the cytoplasm, followed by an NP-40/high salt extraction of the remaining cellular material in HeLa and K4, showed a consistent pattern. Strong GCN1 signal remained in the NP-40/high salt fraction in which membranous and nuclear fractions of the cell are extracted (**Fig 3b**). A traditional Dignam extraction also showed a fraction of GCN1 in the high salt extract after hypotonic extraction (**Fig 3b**). Cell fractionation shows that GCN1 is present in fractions associated with nuclear proteins, but cannot tell us whether GCN1 is present in the nucleus. GCN1 localization within the cell will be further analyzed by immunofluorescence later in the chapter. These results suggest that GCN1 is present in multiple locations within the cell.

#### **HF treatment shifts GCN1 localization**

To determine if the AAR regulates GCN1 localization we tested if HF treatment changes the distribution of GCN1 within the cell. K4 cells were treated with and without HF and sequentially extracted; first by digitonin to selectively permeabilize the cell membrane extracting the cytoplasm, then by NP-40 to extract membranous cell components including the ER, then by high salt to extract nuclear components, and then finally by SDS-containing RIPA buffer to solubilize most of the remaining material. HF increased GCN1 levels in the digitonin fraction, with a concomitant decrease in GCN1 levels in the NP-40 and high salt fractions (**Fig 4**). This indicates that the AAR regulates GCN1 localization in the mammalian cell.

#### **Immunofluorescence does not distinguish a nuclear GCN1 fraction**

To determine whether GCN1 is present in the nucleus we characterized GCN1 localization in the mammalian cell by immunofluorescent staining of GCN1 by confocal microscopy. Staining of both HeLa and K4 cells with Novus GCN1 commercial antibody (same antibody used in immunoblotting) exhibited a nuclear signal, which was clearly visible after digitonin extraction, but was not decreased in shRNA-mediated GCN1 knockdown HeLa (**Fig 5a**). This suggested the nuclear signal was either non-specific or resistant to shRNA knockdown. Staining with a second commercial antibody (Sigma) exhibited no nuclear signal in either cell type and was completely absent in GCN1 knockdown HeLa (**Fig 5b**). It is possible these antibodies recognize different forms of GCN1 since they recognize different parts of GCN1. Comparison of these two commercial GCN1 antibodies by western showed different patterns across biochemical fractions (**Fig S3a**). Since we were unable to generate GCN1 knockout HeLa or K4 cells, which would have provided the definitive control to determine signal specificity, live-cell imaging of cells overexpressing a tdTomato-GCN1 fusion construct was performed to determine which antibody pattern was replicated. tdTomato-GCN1 expression was cytoplasmic (**Fig S3b**). Immunofluorescent staining of transiently expressed Flag-tagged GCN1 exhibited a similar pattern (not shown). Tagged GCN1 and endogenous GCN1 had similar distribution patterns across cellular fractions by immunoblot analysis suggesting neither tag significantly altered GCN1 distribution (**Fig S3a**). Finally since we were able to generate GCN1 knockout HCT116 cells, comparison of immunofluorescent staining by the two commercial GCN1 antibodies in these cells confirmed that the nuclear signal of the first antibody was likely non-specific. The nuclear signal when stained by Novus antibody was present in GCN1 knockout HCT116 cells (**Fig 5c**). There could be differences in different cell types but in combination with the previous data this seems unlikely, supporting cytoplasmic localization of GCN1. No change could be seen in HeLa or K4 cells treated with HF by immunofluorescent staining with either antibody (**Fig 6**).

### **Nuclear IP-MS identifies putative GCN1 interactors and reveals non-specificity of commercial antibody**

To determine what GCN1 may be interacting with in the nucleus, we pulled down GCN1 with commercial antibody from HeLa nuclear lysate in collaboration with the Robin Reed lab, and identified interacting proteins by mass spectrometry. The Reed lab has extensive experience isolating intact and functional endogenous nuclear complexes without overexpression, and have extensive control datasets for comparison to filter out non-specific background. 143 proteins were unique to our GCN1 IP and the 18 most abundant proteins are shown in **Table 1**. There was little overlap with previously documented GCN1 interactors, but 2 mediator subunits were present, albeit only single peptides were detected (not shown). Two interesting groups of interactors emerged including an NFkB1 complex and a Zeb1 complex. In light of GCN1's role in FLS inflammatory gene expression presented in this chapter and known therapeutic activities of HF, the NFkB1 and Zeb1 complexes were particularly interesting as nuclear regulators of inflammation and ECM remodeling. NFkB1 is a major regulator of immune response that forms a complex with MAP3K8 and ABIN-2, which were also pulled down by GCN1 immunoprecipitation (43). Zeb1 is a transcriptional repressor that forms a complex with Ctbp1 and Ctbp2, which were also pulled down (44). The Zeb1 and NFkB1 results were confirmed by pull down of GCN1 with the same commercial GCN1 antibody from HeLa nuclear lysates with subsequent immunoblotting for Zeb1 and NFkB1 (**Fig 7a**). However, immunoprecipitation of Zeb1 and NFkB1 with commercial antibodies was unable to co-precipitated GCN1. Likewise pull down of GCN1 with GCN1 sera generated by our lab was unable to co-precipitate Zeb1 or NFkB1 (Zeb1: not shown, NFkB1:**Fig 7b**). This suggests that these antibodies either impede the respective interaction or commercial GCN1 antibody directly recognizes and pulls down Zeb1 and NFkB1. Ectopic expression of Zeb1 with subsequent immunoblot analysis of the cell lysate with commercial GCN1 antibody shows that the commercial GCN1 antibody directly recognized Zeb1, suggesting that the commercial GCN1 antibody is non-specific (**Fig 7c**). Together with the immunofluorescence results suggesting this antibody has non-specific signal, we do not know which of these proteins are true GCN1 interactors or artifacts of antibody non-specificity. Additional experiments with optimized conditions will be required to confirm which of these proteins are GCN1 interactors.

## **DISCUSSION**

### **GCN1 is required for GCN2-dependent and -independent HF effects in mammals**

The high degree of evolutionary conservation of the AAR pathway throughout eukaryotes suggests GCN1 and ABCF3 as candidate mediators of the GCN2-independent effects of HF treatment. In functional studies in yeast, these genes map upstream of GCN2 and are required for AAR pathway activation. Here I show that in mammalian cells, as in yeast, GCN1 is required for AAR pathway function. Significantly, I also show that GCN1 is required for key GCN2-independent responses to HF-treatment. These observations describe a novel response to nutrient stress in mammalian cells, in that aaRS-inhibition is producing tissue responses that require GCN1 but that do not require the recognized sensor of uncharged tRNA and the signature AAR pathway effector GCN2. These data establish GCN1 as a branch point in the mammalian AAR pathway.

### **ABCF3 is not required for AAR pathway signaling**

Unexpectedly, GCN2 and eIF2 $\alpha$  phosphorylation were not absent or decreased in ABCF3 knockout cells indicating ABCF3 is not required for AAR pathway induction in mammals. Additional experiments also suggested ABCF3 was not required for GCN2-independent GCN1 effects (Kim et al, manuscript in preparation). In yeast, GCN20 is required for full AAR pathway activation: deletion of GCN20 does not eliminate AAR pathway activation, but impairs GCN4 de-repression (yeast ATF4 equivalent) and reduces phospho-eIF2 $\alpha$  levels (22). Although its functional mechanism is not fully understood, in yeast and in *C. elegans*, GCN20/ABCF3 physically interacts with GCN1 via a conserved N-terminal domain (22, 28). In yeast, data suggest that GCN20 stabilizes the interaction between GCN1 and GCN2, and is required for

enhanced GCN1 association with polysomes by ATP (23, 24). Since mammals contain three ABCF family members, compared to two family members in yeast, it is possible that another mammalian ABCF family member or other protein compensates for lost ABCF3 function (45-47). However, ABCF3 is the only mammalian ABCF family member with an N-terminus similar to yeast GCN20 and *C. elegans* ABCF3, suggesting mammalian ABCF3 is a GCN20 ortholog. Further experiments will be required to understand the function of ABCF3, if any, in amino acid sensing in mammalian cells.

### **Functions and Mechanisms of the GCN2-independent GCN1 branch**

GCN1 mediates several GCN2-independent therapeutic transcriptional effects of HF, but whether GCN1 mediates additional GCN2-independent functions or acts outside the inflammatory context is unknown. GCN1 has not been functionally studied in mammalian cells, and its localization within the cell has not been analyzed, but several observations suggest GCN1 may be present in the nucleus. Although GCN1 is assumed to be cytoplasmic, due to its ascribed translational function in yeast, in both yeast and mammals GCN1 putatively interacts with a large number of nuclear proteins, suggesting either that GCN1 is present in the nucleus or might interact with a subset of nuclear proteins that are retained in the cytoplasm (31-40). In addition, this chapter shows that GCN1 is required for GCN2-independent transcriptional changes, also raising the possibility of GCN1 function in the nucleus. We show that HF treatment shifts GCN1 from a digitonin resistant fraction (consistent with organelle association) into a digitonin released cytoplasmic fraction, suggesting that GCN1 association with distinct fractions is dynamically regulated by HF. Immunofluorescence analysis was unable to resolve where this shift corresponded to within the cell, and did not identify a clear nuclear GCN1 fraction. It is possible GCN1 is not present in the nucleus, but this cannot be definitively concluded from the data shown. A small portion of GCN1 could be present in the nucleus below the detectable threshold of our method, or a portion of GCN1 may move into the nucleus under conditions not tested here. Several observations support GCN1 presence in the nucleus, in addition to putative nuclear GCN1 interactors. GCN2 has recently been reported to be present in nucleoli in several mammalian cell lines with nucleolar levels increasing upon amino acid starvation (48). Since GCN2 phosphorylation upon uncharged tRNA accumulation requires GCN1 association, GCN1 may likewise be present in nucleoli. Recent studies have also reported protein translation in the nucleus, and GCN1 as a translational component, could be associated with ribosomal components in the nucleus (49). In addition, tRNAs accumulate in the nucleus upon nutrient deprivation, which could plausibly play regulatory roles by activating nuclear GCN2 and GCN1 (50).

### **Further experiments are required to confirm GCN1 nuclear interactors**

GCN1 is a large, scaffold-like protein whose amino acid sequence contains numerous protein-protein interaction elements, including LxxLL motifs and HEAT domains, which in other proteins are known to interact with nuclear proteins (51). Several studies have identified GCN1 as an interactor with a large number of nuclear regulators, including transcription factors e.g. SREBP, components of the splicing apparatus e.g. CDC5L, components of the transcriptional machinery: components of the mediator complex, and several nuclear receptors (31-40). My own studies, however, have failed to reproduce several previously identified interactions (not shown). Our analysis of GCN1 interactors from nuclear-associated lysate indicated little overlap with previously identified interactors, although follow-up experiments revealed commercial GCN1 antibody was not sufficiently specific to reliably identify GCN1 co-precipitating proteins. The limitations of the commercial antibody shown here demonstrate that an alternative approach will be required to identify GCN1 protein interactors.

A branch point exists in the mammalian AAR and GCN1 mediates key GCN2-independent therapeutic transcriptional changes upon HF treatment, but how these effects are mediated remains unclear. In the next chapter we return to the system and conditions in which we know the GCN2-independent GCN1 branch mediates effects and characterize a full transcriptional program mediated by this new branch in the presence and absence of inflammatory stimulation. In addition, learning from the experiments in this

chapter, we set up an improved system to pull down GCN1 to identify potential pathway mediators that regulate the identified transcriptional changes.

## METHODS

### Reagents

HF and HFol (Prepared by Dr. R.Mazitschek, MGH)

TNF $\alpha$  (Sigma)

Effectene (Qiagen) was used for transient transfections.

ANTIBODIES: GCN1 (Novus NB100-97851: used for all westerns and IPs unless otherwise indicated, Sigma HPA 019648), pGCN2 (Abcam/Epitomics EPR2320Y /AB75836), GCN2 (CST 3302S), p $\epsilon$ IF2 $\alpha$  (3398S),  $\epsilon$ IF2 $\alpha$  (CST 9722), ABCF3 (Novus NBP1-84781: used for all westerns unless otherwise indicated, Bethyl A302-496A-T), Actin (Sigma A5441), NF $\kappa$ B1 (CST 13586S and 3035S), Zeb1 (Bethyl A301-922A-T), c-Myc (CST 5605, CREB (CST 9197, PARP (CST 9532S), H3 (CST 3638S), S6 (CST 2317), GAPDH (Genetex GTX627408), ERP29 (Genetex GTX102225), GCN1 sera (generated against GCN1 epitope: QFFVPDALNDRHPDVRKC, Abfrontiers)

### Cell Lines and Cell Culture

HCT116 (ATCC), K4 (Dr Evelyn Murphy (University College Dublin)), HeLa (ATCC), and 293T (ATCC) were cultured at 5% CO<sub>2</sub> at 37°C in DMEM supplemented with 10% FBS, L-glutamine, and antibiotics (K4 and HCT116 cells were further supplemented with sodium pyruvate).

For TNF $\alpha$ -stimulated experiments, cells were shifted into K4 cell media with 0.2% FBS instead of 10% FBS 24 hours before the HFol treatment. HFol or carrier alone (DMF) were added directly to cells without media change. After 16 hours, 10 ng/mL TNF $\alpha$  (Sigma) was added directly to cells. After 8 hours, RNA was extracted for analysis.

### cDNAs

The following cDNAs were used: pCSX, pCSX-eGFP, pCSX-3Flag-hGCN1 (N-terminally tagged), tdTomato-C1-hGCN1 (N-terminally tagged) (hGCN1 from pCSX-3Flag-hGCN1 cloned into tdTomato using XbaI and EcoRI), pCSX-3Flag-shResistant-hGCN1 (N-terminally tagged) (shRNA GCN1 target sequence codons were changed to wobble codons to encode same GCN1 amino acid sequence but to make nucleotide sequence different from target sequence of shRNA. shRNA target sequence mutated to ACACAGGCTTACAGGTAATAA using gBlock sequence (IDT) provided below to replace portion of GCN1 in pCSX-3Flag-GCN1 with BstEII and BclII.), pEV-p105-WT (Addgene 23952) NF $\kappa$ B1, pCSX-3Flag-hZeb1 (hZeb1 cDNA cloned from pCR-Blunt II-TOPO MGC human Zeb1 sequence-verified cDNA, clone ID: 40036600 (Dharmacon)).

gBlock sequence:

```
ACTTGGGTACCTGAAATCCAACAGTCCCCGCCTGATGGATGAAGCTGTGCTGGCACTGCGG
AACCTGGCACGCCAGTGCAGTGACTCTTCGGCCATGGAATCCCTGACCAAGCACCTATTTGC
TATCCTCGGAGGCTCGGAAGGAAACTAACTGTTGTAGCCCAGAAGATGAGCGTCCTCTCAG
GGATTGGGAGCGTCAGTCATCACGTGGTGTCTGGACCTTCCAGTCAGGTCCTGAATGGGATC
GTGGCTGAGCTGTTTCATCCCGTTCCTTCAGCAGGAAGTTCATGAAGGGACCTTGGTACACGC
TGCTCAGTCCTGGCTCTCTGGTGTAAACCGATTCACTATGGAAGTGCCCAAGAAGCTCACTG
AATGGTTCAAAAAAGCTTTCAGCCTTAAACCTCCACATCTGCGGTGAGGCATGCCTACCTG
CAGTGCATGTTGGCCTCTTACCGGGGTGACACGCTGTTGCAGGCCCTGGACTTACTGCCCTT
GCTCATCCAGACAGTGGAGAAGGCAGCCTCCCAAAGCACTCAGGTTCCCAACCATCACCGAA
GGGGTTGCCGCAGCCTTGTGTCTTAAAGTTGTGAGTGGCTGACTCACAGGCTGAGGCCAA
ACTGAGCAGTTTCTGGCAGTTGATTGTGGATGAGAAAAAGCAGGTTTTCACTTCTGAGAAAT
TCCTGGTCATGGCTTCAGAGGATGCCCTGTGTACTGTGTTGCATCTGACAGAGAGACTTTTCC
TTGACCACCCACACAGGCTTACAGGTAATAAAGTTCAGCAGTACCACCGGGCTCTGGTGGCG
GTGCTCCTGAGCCGCACCTGGCACGTCCGCAGGCAGGCTCAGCAGACAGTTCCGGAAGCTGCT
GTCCTCTCTTGGGGGCTTTAAGCTGGCGCACGGAAGTCTTGGAGGAGCTGAAGACTGTCCTCA
```

GTTCTCACAAGGTGCTGCCCTTAGAGGCTTTGGTGACTGATGCTGGAGAGGTGACTGAGGCA  
GGCAAGGCCTACGTGCCTCCACGGGTCCTGCAGGAGGCTCTGTGTGTCATCTCCGGTGTGCC  
AGGGCTCAAGGGTGATGTACCGACACTGAACAACTGGCCCAGGAAATGCTGATCAAAGTC

### **CRISPR-mediated knockout HCT116 cells**

HCT116 ABCF3 knockout cells were generated using the Origene knockout kit according to the manufacturer's protocol (Catalogue #: KN201403: contains gRNA containing plasmids KN201403G1 and KN201403G2, and donor cassette containing plasmid KN20143D) Clonal lines were isolated using cloning rings (Sigma). For genomic PCR confirmation, DNA was extracted using DNeasy Blood and Tissue Kit (Qiagen). PCR was performed using ChoiceTaq Polymerase (Denville Scientific) according to manufacturer's protocol. To prevent contamination, dTTP was replaced with dUTP, and 1 unit Uracil-DNA Glycosylase (UNG) was added to PCR reactions, and incubated at RT for 10 minutes prior to PCR.

ABCF3 Guide RNA Target Sequences:

G#1: GCCGATTCACTTCTGCTTAC, G#2: CAGAACTGAATCGGCCGGCT,

gSc: GCACTACCAGAGCTAACTCA

Genomic PCR Primer Sequences:

For left side homology template: lha42f: TGCAAGGAAGGGCGAGATTT,

template242r: CTCGTCGCTCTCCATGCTAG

For right side homology template: template2657F: ATCAGCAGCCTCTGTTCAC,

rha3037R: GTTCCCCTACAGCTTCCACC

For genome: genome42F: TGCAAGGAAGGGCGAGATTT,

genome498R: GTTCCCCTACAGCTTCCACC

GCN1 and GCN2 CRISPR-mediated knockout HCT116 cell lines were generated and obtained from Changqian Zhao.

### **shRNA-mediated knockdown GCN1 and GCN2 HeLa and K4 by lentiviral infection**

Virus was produced by co-transfecting the transfer vector, pLKO.1 containing the shRNA inserts given below, with the envelope plasmid pMD2.G, and packaging plasmid psPAX2, into 293T cells plated on poly-D-lysine coated plates at a ratio of 2:1:1.5, respectively. Media was changed 24 hours after transfection and the virus-containing supernatant was collected 72 hours after transfection and passed through a 0.45 µm filter to eliminate cells. Target HeLa or K4 cells in 6-well tissue culture plates were infected in media containing 5 µg/mL of polybrene. Media was changed after 6 hours. After 16 hours infection was repeated. After 24 hours, cells were selected with puromycin. After selection, cells were maintained in respective cell media with 0.25 µg/mL puromycin.

pLKO.1 shGCN1: HsSH00265698 (Dana Farber/Harvard Cancer Center DNA Resource Core)

**CCGGGCATAGACTCACTGGCAACA**CTCGAGTTGTTGCCAGTGAGTCTATGCTTTTT (target sequence in bold)

pLKO.1 shGCN2: HsSH00114723 (Dana Farber/Harvard Cancer Center DNA Resource Core)

**CCGGGCCTAACTGGTGAAGAAGTATCTCGAGATACTTCTTACCAGTTAGGCTTTTTG**  
(target sequence in bold)

pLKO.1 shScrambled: 1864 (Addgene)

**CCTAAGGTTAAGTCGCCCTCGCTCGAGCGAGGGCGACTTAACCTTAGG**

shRNA-mediated scrambled, GCN1, and GCN2 knockdown K4 cells were generated and obtained from Yeon-Jin Kim.

### **Cell Fractionations**

**KIT:** HeLa and K4 cells were fractionated with the Thermo Subcellular Protein Fractionation Kit for Cultured Cells (78840) according to the manufacturer's protocol with the following modifications. Cells were fractionated from 10 cm tissue culture plates and reagent volumes were based on a packed cell volume of 50  $\mu$ L. After cytoplasmic extractions (step 2) and membrane extractions (step 5), pellets were washed twice with PBS followed by 2 minute 500G spins, or twice with MEB followed by 3000 G spins, respectively. In step 10, the extraction was incubated for 10 min at 37C and briefly vortexed after 5 minutes.

**DIGITONIN SEQUENTIAL EXTRACTION:** Cells were sequentially permeabilized with digitonin buffer (110 mM KOAc, 25 mM KHEPES pH 7.2, 2.5 mM MgCl<sub>2</sub>, 1 mM EGTA, 0.015% digitonin) for 5 minutes, washed with wash buffer D (110 mM KOAc, 25 mM KHEPES pH 7.2, 2.5 mM MgCl<sub>2</sub>, 1 mM EGTA, 0.004% digitonin) for 1 minute, lysed with NP-40 buffer (110 mM KOAc, 25 mM KHEPES pH 7.2, 2.5 mM MgCl<sub>2</sub>, 1% NP-40) for 5 minutes, washed with wash buffer N (110 mM KOAc, 25 mM KHEPES, pH 7.2, 2.5 mM MgCl<sub>2</sub>, 1% NP-40) for 1 minute, and lysed with High Salt buffer (400 mM KOAc, 25 mM KHEPES, pH 7.2, 15 mM MgCl<sub>2</sub>, 1% NP-40), washed with wash buffer H (400 mM KOAc, 25 mM KHEPES, pH 7.2, 15 mM MgCl<sub>2</sub>, 1% NP-40), and lysed with RIPA. Protease and phosphatase inhibitors were also added to all lysis buffers. In figure 3 the following modifications were performed: cells were scraped into and lysed in microcentrifuge tubes rather than lysed directly on the plate and spun between subsequent extractions, and the NP-40/high salt extraction was performed together and incubated for 30 minutes rather than 5 minutes.

**HYPOTONIC/HIGH SALT DIGNAM LYSIS:** Based on Folco et al. 2012 (52). Cells were resuspended in hypotonic buffer (10mM HEPES, pH 7.9, 1.5mM MgCl<sub>2</sub>, 10mM KCl, 0.2mM PMSF, 0.5mM DTT) and spun. The pellet was resuspended in 3 volumes hypotonic buffer and incubated for 10 minutes on ice. Cells were briefly homogenized with a mini-dounce and spun. The supernatant was saved as the hypotonic fraction. The pellet was further extracted in high salt buffer (20mM HEPES, pH 7.9, 1.5mM MgCl<sub>2</sub>, 0.8 M KCl, 0.2mM EDTA, 25% Glycerol, 0.2mM PMSF, 0.5mM DTT) according to the reference, and spun. The supernatant was saved as the high salt fraction.

### **Protein Extraction, SDS PAGE, and Immunoblotting**

Protein content was quantified with the Pierce BCA protein assay kit (Thermo Scientific) except for cell fractionations. Protein samples were diluted in 3X Laemmli buffer (160 mM Tris pH 6.8, 6% SDS, 20% glycerol, 0.01% bromophenol blue, 0.04%  $\beta$ -Mercaptoethanol), boiled for 5 minutes, and electrophoresed through pre-cast Bolt 4-12 % Bis-Tris Plus gels (Invitrogen). The amount of total protein loaded per well was standardized across each experiment except for cell fractionation experiments. Proteins were transferred to nitrocellulose membrane (GE Healthcare), blocked with 5% milk and blotted with appropriate antibodies overnight at 4°C. Blot were imaged with the Syngene PXi imager.

### **RNA Extraction and Quantitative Real-Time PCR**

Cells were extracted with 1 mL TRIZOL (Ambion) and RNA purified with the Direct-Zol RNA MiniPrep kit (Zymo Research). RNA yield and purity were quantified using an ND-1000 spectrophotometer and cDNA was synthesized using the ABI High Capacity cDNA Reverse Transcription Kit (Ambion). Primer pairs were designed using Universal ProbeLibrary Assay Design Center (<https://qpcr.probefinder.com>). The following primer pairs and probes were used: TRIB3 (TCCAGATCGTGCAACTGCT, CTCCTGGACGGGGTACA, #16), GAPDH (AGCCACATCGCTCAGACAC, GCCAATACGACCAAATCC, #60), ASNS (GATGAACTTACGCAGGGTTACA, CGCGGAGAACATCAAACAA, #63), ATF3 (TTTGCCATCCAGAACAAGC, CATCTTCTTCAGGGGCTACCT, #53) CXCL10 (GAAAGCAGTTAGCAAGGAAAGGT, GACATATACTCCATGTAGGGAAGTGA, #34), MMP13 (CCAGTCTCCGAGGAGAAACA, AAAACAGCTCCGCATCAAC, #73), IDO1 (GTGTTTCACCAAATCCACGAT,



CTGATAGCTGGGGGTTGC, #20). Quantitative real-time PCR was performed using the LightCycler 480 ProbesMaster and UPR Light Cycler System (Roche). Relative gene expression changes were measured using the comparative  $C_T$  method,  $X=2^{-\Delta\Delta C_T}$ . Target gene expression was normalized with GAPDH as the respective endogenous control.

### **Immunofluorescence and Microscopy**

Cells were plated on sterilized coverslips (#1.5 12 mm round, Warner CS-12R17) in 24-well dishes. Cells were washed once with PBS and fixed with 500  $\mu$ L 2% paraformaldehyde in 1X PBS for 30 minutes. Fixed cells were washed twice with 500  $\mu$ L PBS for 5 minutes. Cells were permeabilized with 0.25% Triton-X-100 in 10% NGS, 1X PBS for 30 minutes. Cells were incubated overnight at 4°C with primary antibody diluted in 0.1% Triton in 10% NGS, 1X PBS at 1:20 or 1:50. Cells were washed 3 times with PN buffer (1% NGS, 1X PBS), 1X with 1M LiCl, and 2X with PN buffer for 5 minutes each. Cells were then incubated in anti-rabbit Alexa488 conjugated secondary antibody (CST) diluted at 1:500 in 0.1% Triton-x-100 in 10% NGS, 1X PBS for 1 hour. Cells were further washed 1X with 0.1% Triton-x-100 PN for 5 minutes, incubated 10 minutes with 0.5  $\mu$ g/mL Hoeschst 33258 to stain nuclei, followed by 2X 0.1% Triton-x-100 PN for 10 minutes each, and 3X with PN for 10 minutes each. All steps were performed at room temperature unless indicated otherwise. Cover slips were mounted cell side down and imaged with a Yokogawa CSU-X1 spinning disk confocal on a Nikon Ti fluorescence microscope with the 100X oil Plan Apo  $\lambda$  N.A. 1.45 Type: DIC N2 objective, at the Nikon Imaging Center, Harvard Medical School. Images were analyzed using ImageJ.

### **IP-MS**

HeLa nuclear extracts for the IP-MS were prepared and immuno-precipitated by Binkai Chi in the Robin Reed lab, Harvard Medical School. Nuclear lysates were prepared according to Folco et al. 2012 (52). Novus GCN1 antibody (NB100-97851) was crosslinked to Protein A beads in a ratio of 2:1 antibody:beads with dimethyl pimelimidate. Prior to immunoprecipitation 75  $\mu$ L of HeLa nuclear lysate was incubated at 30°C for 20 min with 75  $\mu$ L SDB (20 mM HEPES pH 7.9, 100 mM KCl), 10  $\mu$ L 12.5 mM ATP, 10  $\mu$ L 80 mM MgCl<sub>2</sub>, 10  $\mu$ L 0.5M creatine phosphate, and 70  $\mu$ L water. After incubation, mix was spun at 14,000 rpm at 4°C. 15  $\mu$ L was taken as an input control sample. 125  $\mu$ L IP buffer (1x PBS, 0.1% Triton-X100, 0.6 mM PMSF, 2x 25X Roche Complete Mini, EDTA-free protease inhibitor) was added to remaining mix, combined with 20  $\mu$ L GCN1-crosslinked protein A beads, and incubated at 4°C for 4h or overnight. Beads were washed 6x with 0.1% Triton-X-100, 0.2 mM PMSF, 1x with PBS, and eluted with 60  $\mu$ L PGB (138 mM Tris, 4.5% SDS, 22% glycerol, bromophenol blue) for 20 min at room temperature. Eluate was TCA precipitated and sent for LC/MS/MS analysis at the Taplin Mass Spectrometry Facility, Harvard Medical School. Identified proteins were compared to two control IP-MS sample sets from the Reed lab (matrin IP and IgG IP).

### **Follow-up NFkB1 and Zeb1 IPs**

Confirmation of NFkB1 and Zeb1 co-immunoprecipitation with Novus GCN1 antibody was performed as above with the following modifications: GCN1 antibody was coupled rather than crosslinked to protein A beads for 1 hour at room temperature, eluate was analyzed by immunoblot, and immunoprecipitation with equivalent amount of IgG was used as control IP.

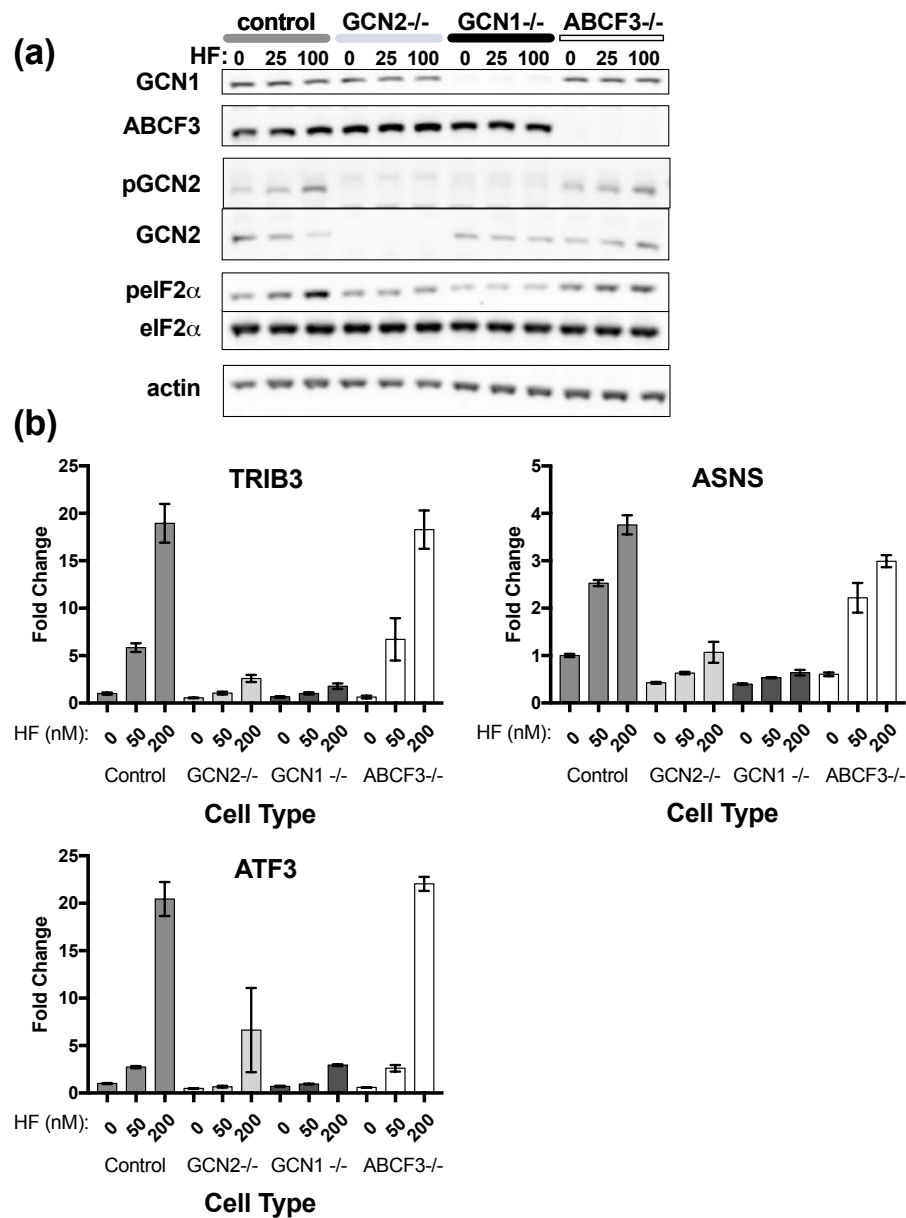
Follow up coIPs in Figure 7b were immunoprecipitated with respective antibodies or IgG from HeLa cells lysed with 1% triton (20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, 1% triton + protease inhibitors (Complete and PMSF). Lysate was incubated with primary antibody at recommended concentrations overnight at 4°C. Protein A beads were added and incubated for an additional 4 hours, washed 6x with lysis buffer, eluted in Laemmli buffer and analyzed by immunoblot. For GCN1 sera IP the following modification were made. Sera was coupled to beads at a 5:1 ratio of sera:beads for 1 hour at room temperature. A fraction of coupled beads were incubated with GCN1 sera peptide as control.

Sera coupled beads and peptide blocked sera coupled beads were then incubated with cell lysate overnight.

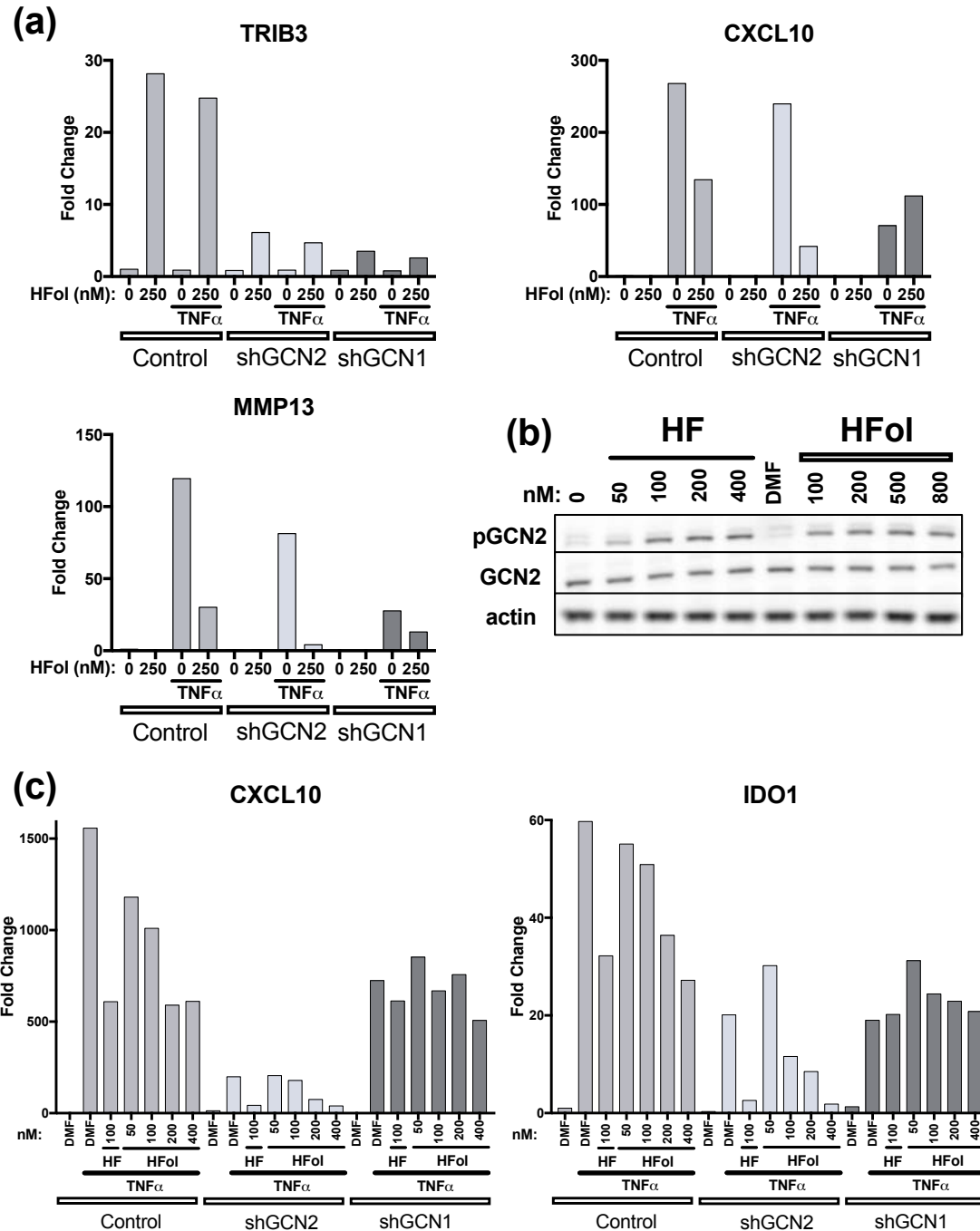
**Statistics**

Data were analyzed using GraphPad PRISM, version 6.

## FIGURES AND TABLES

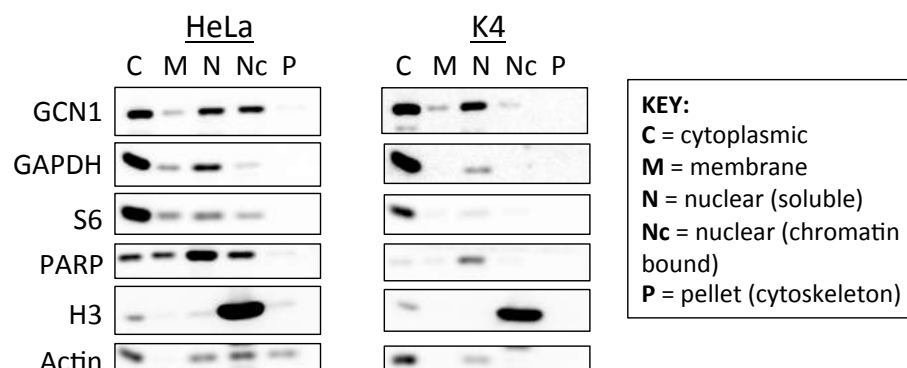


**Fig 1 | GCN1, but not ABCF3 is required for phosphorylation of GCN2 by the AAR in mammalian cells.** **(a)** GCN1, GCN2, and ABCF3 knockout HCT116 cell lines were treated with 25 or 100 nM HF for 30 minutes and assayed by immunoblot for phospho-GCN2 and phospho-eIF2 $\alpha$ . Total GCN1, ABCF3, GCN2, eIF2 $\alpha$ , and actin are shown as loading controls and to demonstrate successful knockout in respective knockout cell line. **(b)** mRNA expression of characteristic AAR induced genes from GCN1, GCN2, and ABCF3 knockout HCT116 cell lines treated with 50 or 200 nM HF for 8 hours. mRNA expression was normalized to expression of GAPDH mRNA and is represented as the mean of biological triplicates  $\pm$  s.d. GCN2<sup>-/-</sup> and GCN1<sup>-/-</sup> HCT116 data in (a) are representative of a single experiment but mirror observations observed by shRNA-mediated knockdown of GCN2 and GCN1 in multiple cell types (not shown). ABCF3<sup>-/-</sup> HCT116 data in (a) are representative of 3 separate experiments. Data in (b) are representative of a single experiment but mirror observations observed by shRNA-mediated knockdown of GCN2, GCN1, and ABCF3 in K4 cells (not shown, Kim et al, manuscript in preparation).

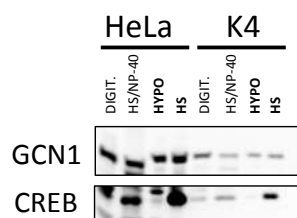


**Fig 2 | Several key HF therapeutic effects require GCN1 but not GCN2.** (a) mRNA expression of TRIB3 (characteristic AAR pathway induced gene) and CXCL10 and MMP13 (proinflammatory genes) in shRNA-mediated knockdown control (shScrambled), GCN1 (shGCN1), and GCN2 (shGCN2) K4 cells treated with 250 nM HFol for 24 hours and 10 ng/mL TNFα for 8 hours. (b) Comparison of GCN2 phosphorylation by immunoblot of wildtype K4 treated with HF or HFol at indicated concentrations for 30 minutes. (c) Comparison of CXCL10 and IDO1 mRNA expression (proinflammatory genes) with HF or HFol treatment at indicated concentrations and TNFα treatment in shRNA-mediated knockdown K4s as described in (a). Data in (a) and (c) are representative of 6 independent experiments. Data in (b) is representative of 1 experiment but has been repeated in other cell types.

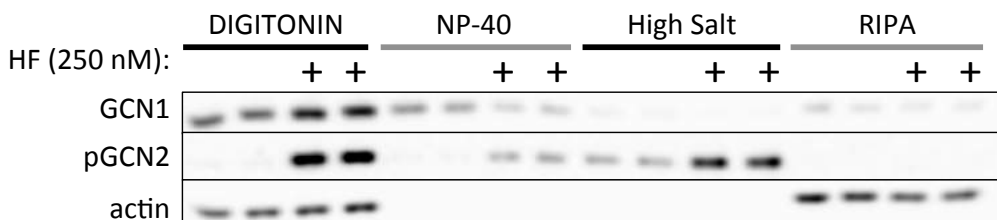
**(a) Pierce Commercial Extraction kit:**



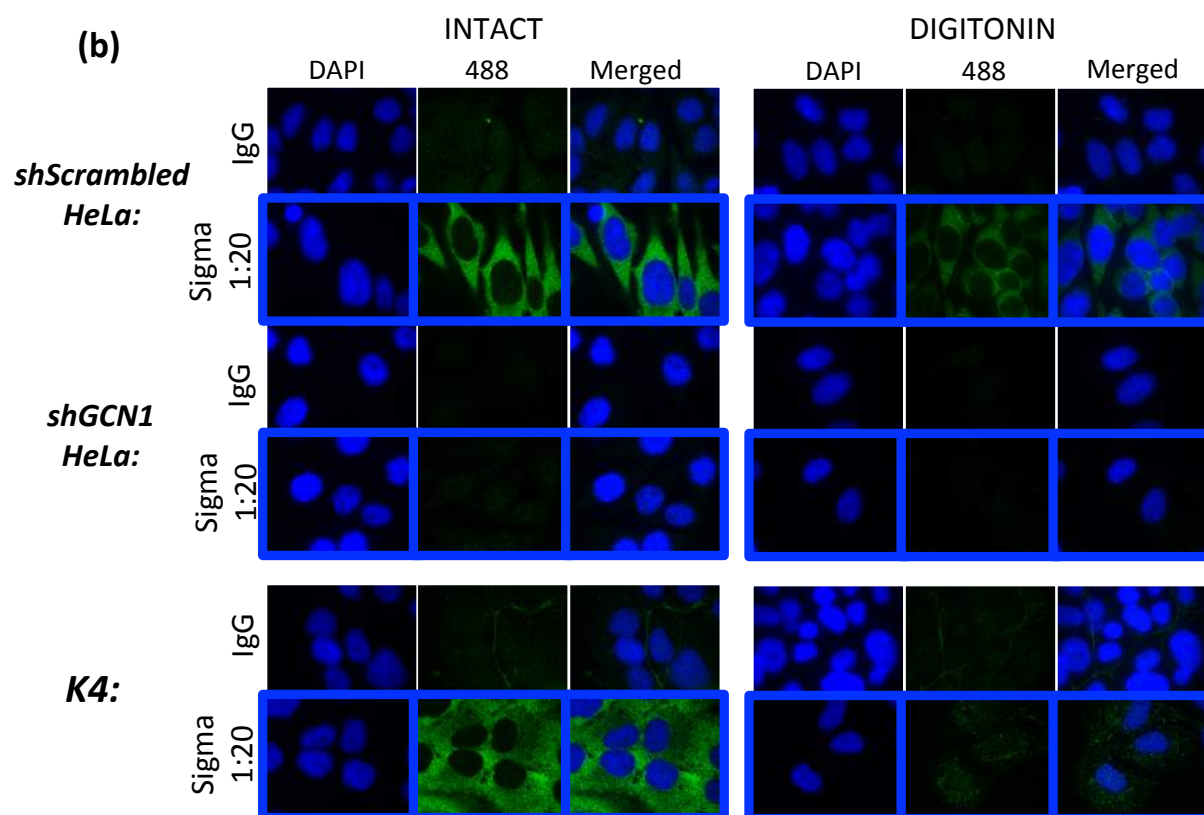
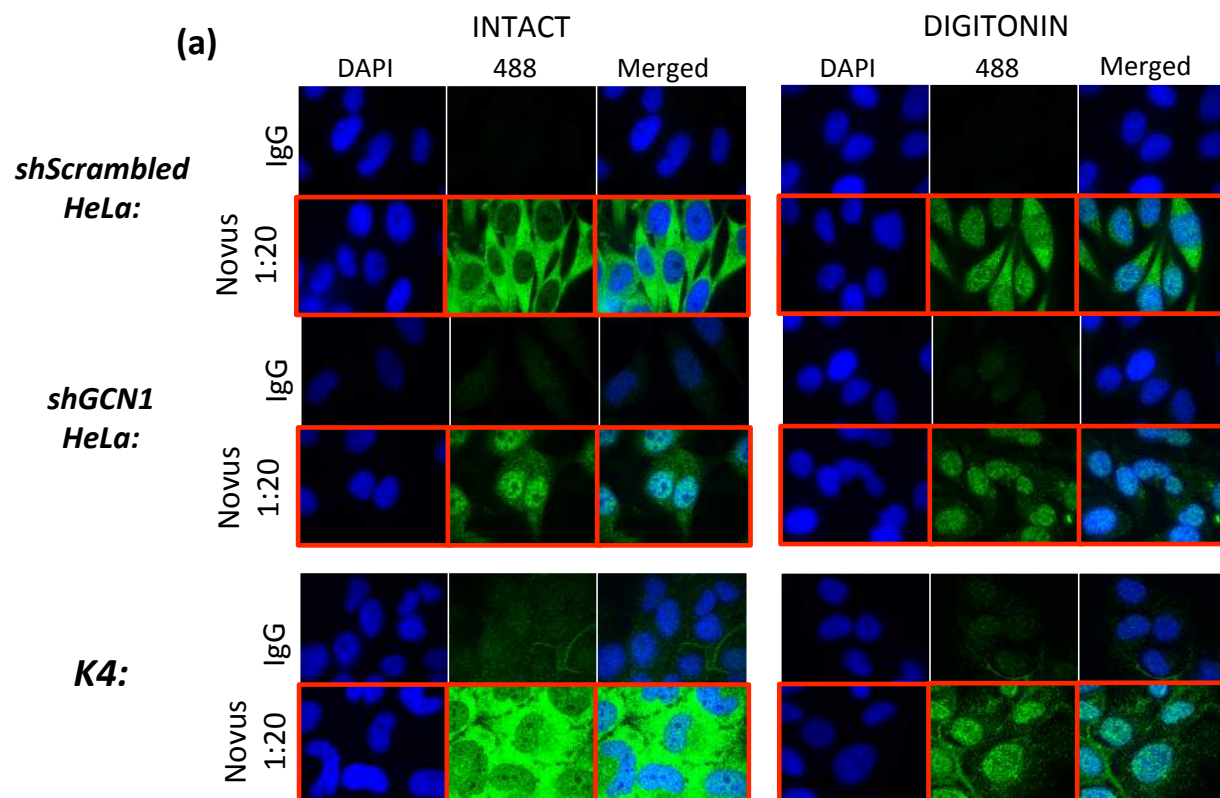
**(b) Digitonin Sequential Extraction and Dignam Extraction:**

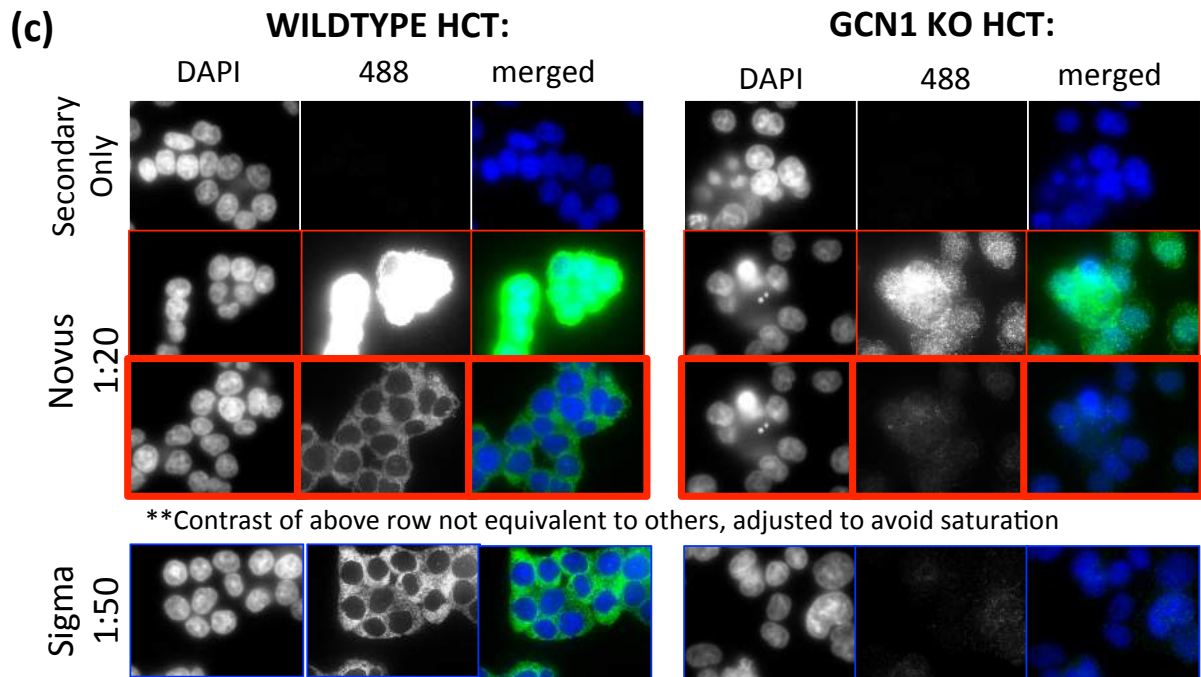


**Fig 3 | Several different biochemical fractionation methods consistently show presence of GCN1 in cell fractions that contain nuclear components.** (a) Fractionation of HeLa and K4 cells using the Thermo Subcellular Protein Fractionation Kit for Cultured Cells and analysis by immunoblot shows a significant amount of GCN1 is present in the nuclear-associated fractions. GAPDH, S6, PARP, H3, and actin are present in specific fractions as expected. (b) Digitonin sequential extraction and Dignam extraction compared in HeLa and K4 cells. Digitonin sequential extraction consisted of an initial extraction using digitonin, to selectively extract the cytoplasm, followed by extraction of the remaining pellet with NP-40 and high salt. Dignam extraction consisted of hypotonic lysis followed by high salt extraction of the remaining pellet. HeLa data in (a) is representative of 2 separate experiments, while K4 data is representative of a single experiment. Similar results were obtained in both cell lines using additional commercial extraction kits (data not shown). Data in (b) are representative of at least 3 separate experiments.

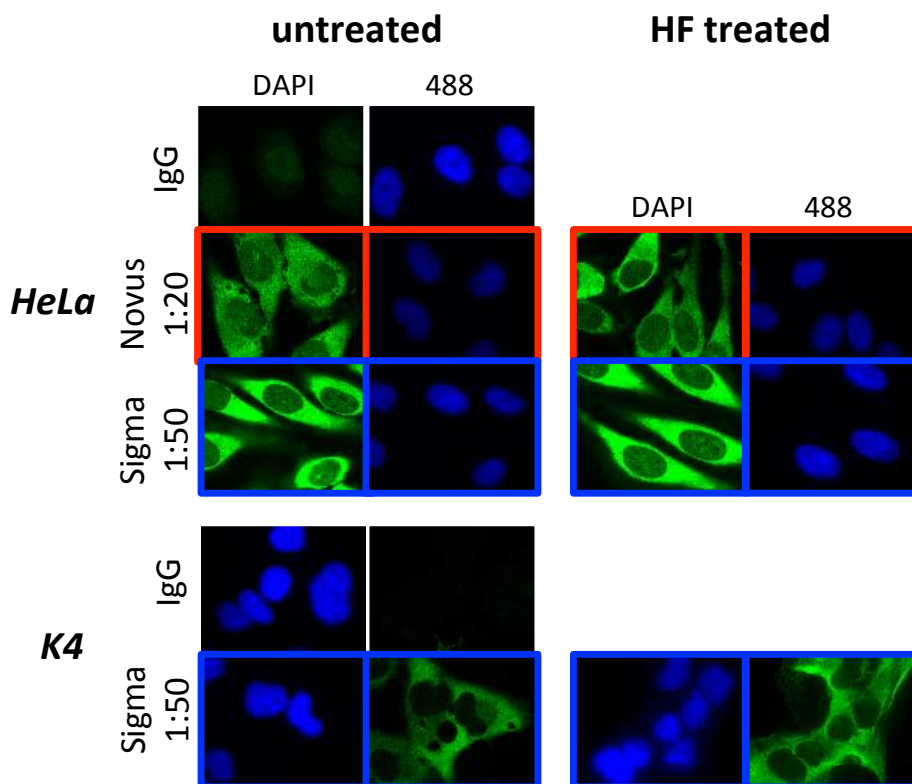


**Fig 4 | HF treatment shifts GCN1 localization between cellular fractions.** Sequential digitonin extraction of K4 cells treated with 250 nM HF for 30 minutes. Cells were sequentially extracted on 10 cm plates first by digitonin, then NP-40, then high salt, and finally all remaining material was extracted with RIPA. Data are representative of at least 3 independent experiments.





**Fig 5 | GCN1 localizes to cytoplasm in multiple mammalian cells.** (a) Images of HeLa, shRNA-mediated GCN1 knockdown HeLa, and K4 cells stained with commercial Novus GCN1 antibody with and without digitonin permeabilization. (b) Same cells as (a) stained with commercial Sigma GCN1 antibody with and without digitonin permeabilization. (c) Images of wildtype and GCN1 knockout HCT116 cells stained with Novus or Sigma GCN1 antibodies. Third row is identical sample as second row with adjusted contrast to avoid saturation. For all images antibody concentrations are indicated and IgG stained samples are provided as controls. Data in (a) and (b) are representative of at least 3 separate experiments. Data in (c) are representative of a single experiment.

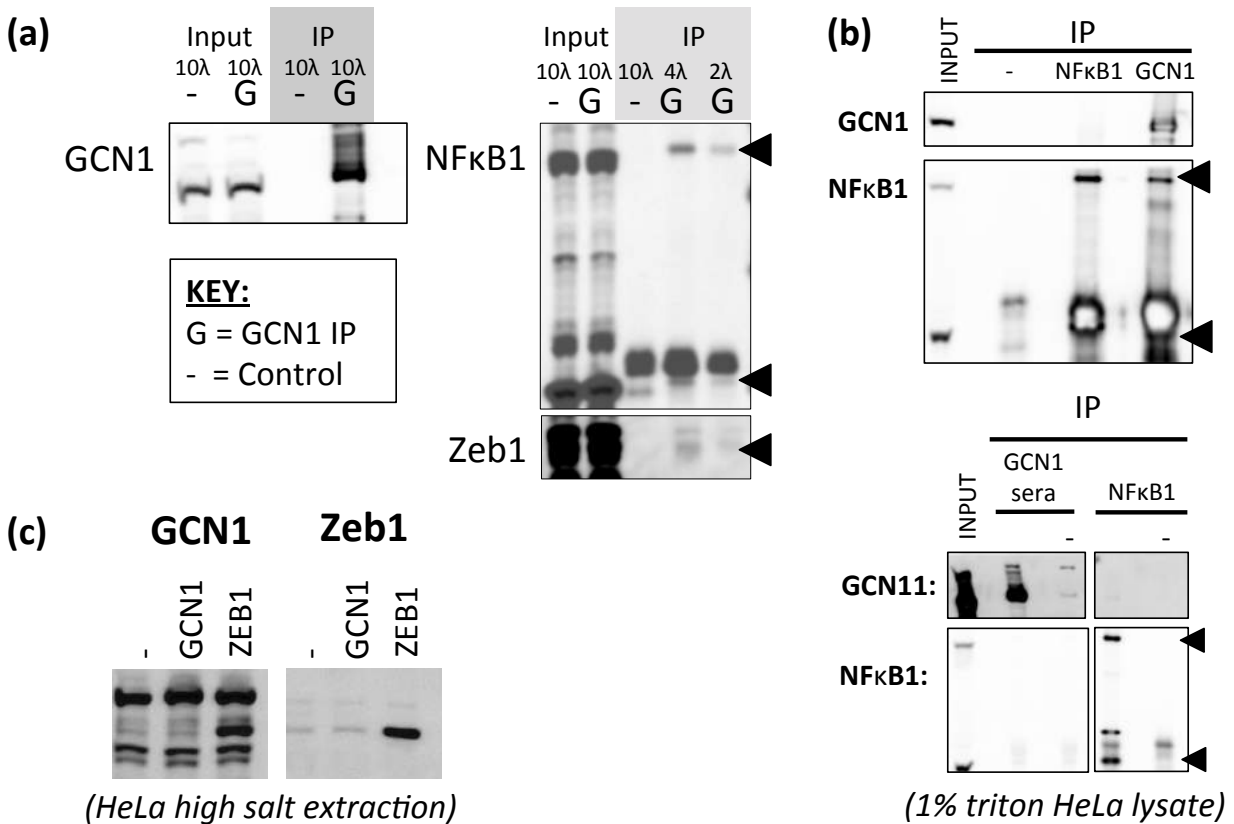


**Fig 6 | Shift in GCN1 localization by HF treatment is not observed by immunofluorescence.** (top) Images of HeLa stained with Novus and Sigma GCN1 antibodies with and without treatment with 800 nM HF for 30 minutes. (bottom) Images of K4 stained with Sigma GCN1 antibody with and without treatment with 400 nM HF for 3 hours. For all images antibody concentrations are indicated and IgG stained samples are provided as controls. Data are representative of at least 2 separate experiments.



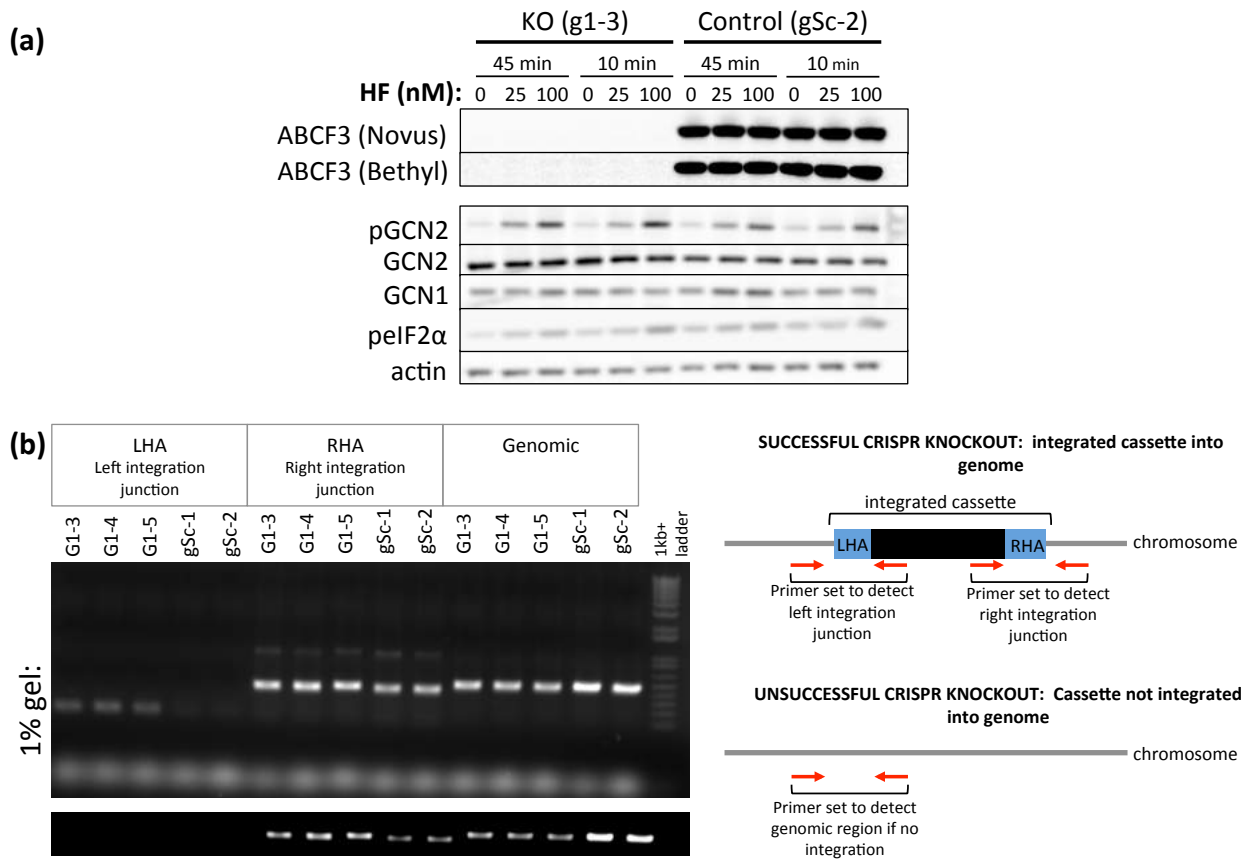
Unique:	Total:	reference:	Gene Symbol:	MW (kDa):
189	421	Q92616_GCN1L_HUMAN	GCN1L1	292.57
24	25	P37275_ZEB1_HUMAN	ZEB1	124
22	22	Q9ULJ3_ZBT21_HUMAN	ZBTB21	118.8
20	20	P19838_NFKB1_HUMAN	NFKB1	105.29
18	19	Q7Z6Z7_HUWE1_HUMAN	HUWE1	481.59
9	11	Q96SU4_OSBL9_HUMAN	OSBPL9	83.13
9	9	Q14202_ZMYM3_HUMAN	ZMYM3	152.28
8	9	P34931_HS71L_HUMAN	HSPA1L	70.33
8	9	O15523_DDX3Y_HUMAN	DDX3Y	73.11
8	8	Q92734_TFG_HUMAN	TFG	43.42
8	8	O75145_LIPA3_HUMAN	PPFIA3	133.41
7	7	P56545_CTBP2_HUMAN	CTBP2	48.91
7	7	Q13363_CTBP1_HUMAN	CTBP1	47.51
6	6	P20936_RASA1_HUMAN	RASA1	116.33
6	6	P54652_HSP72_HUMAN	HSPA2	69.98
6	6	P00338_LDHA_HUMAN	LDHA	36.67
5	5	P17980_PRS6A_HUMAN	PSMC3	49.17
5	5	Q8NFZ5_TNIP2_HUMAN	TNIP2	48.67

**Table 1 | Nuclear isolation of GCN1 from HeLa cells revealed putative interactors but also non-specificity of commercial antibody.** GCN1 was immunoprecipitated from HeLa nuclear extract with commercial antibody (Novus), eluted, and analyzed by mass spectrometry to identify nuclear GCN1 interactors. Table shows 18 most abundant proteins found in the eluate.

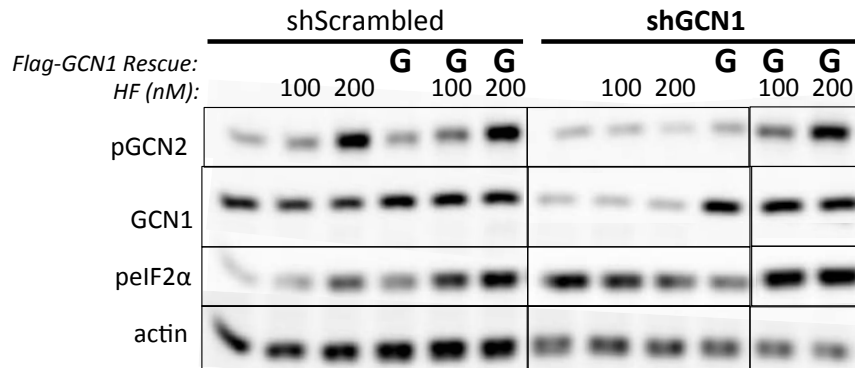


**Fig 7 | Follow up experiments indicate commercial GCN1 antibody has non-specific target recognition.** (a) IP-MS NFkB1 and Zeb1 results are repeatable when HeLa high salt lysate is immunoprecipitated with commercial GCN1 antibody (Novus). Arrows indicate p50 and p105 forms of NFkB1, and Zeb1 band. Data is representative of two separate experiments. (b) Antibody pull downs of GCN1 and NFkB1 from HeLa 1% triton lysates. (top panel) Commercial GCN1 (Novus) antibody co-precipitates NFkB1, but NFkB1 pull down does not co-precipitate GCN1. (bottom panel) GCN1 sera recognizing a distinct epitope from the commercial GCN1 (Novus) antibody does not co-precipitate NFkB1. A second NFkB1 antibody recognizing a different epitope from the antibody used in top panel does not co-precipitate GCN1. (c) Immunoblot of HeLa high salt extract with ectopically expressed GCN1 or Zeb1 shows that commercial GCN1 antibody (Novus) recognizes a ZEB1 transfection dependent band at exact size of Zeb1.

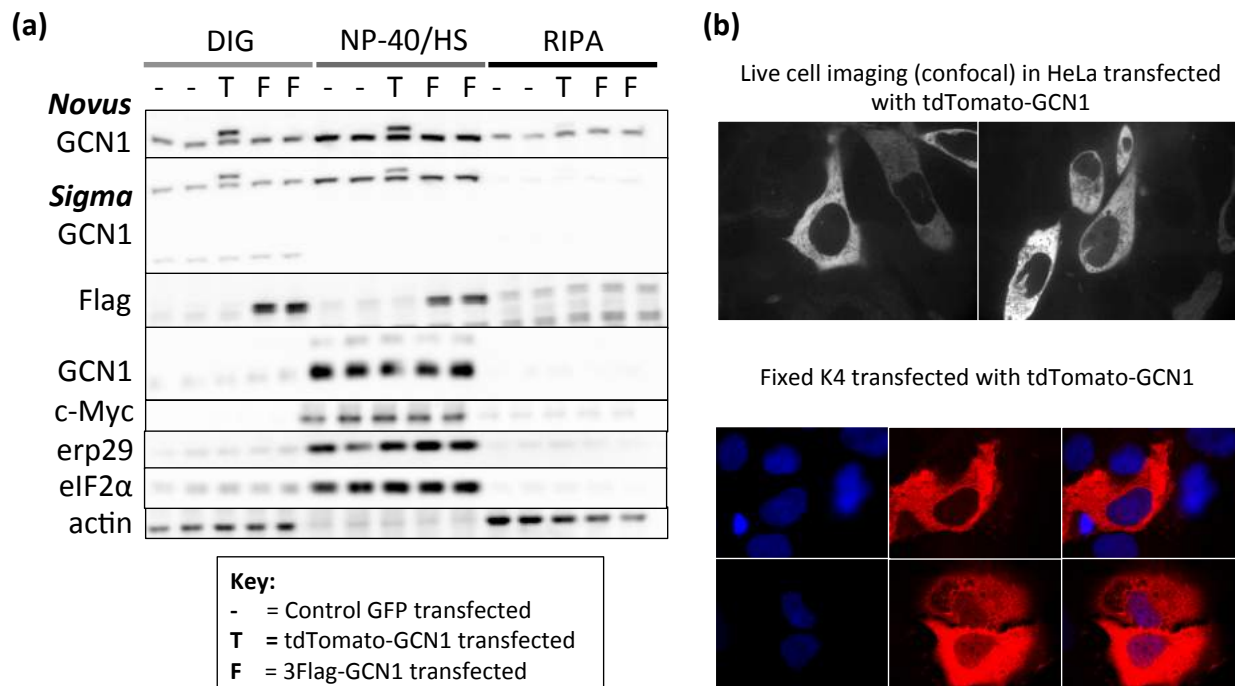
SUPPLEMENTARY FIGURES



**Fig S1 | Confirmation of ABCF3 knockout HCT116 cells.** **(a)** Comparison of ABCF3 protein levels with two different commercial antibodies and AAR pathway activation by HF treatment by immunoblot analysis. **(b)** Genomic PCR confirmation of ABCF3 knockout clonal lines. Gel showing PCR fragments resulting from PCR reactions using primers described by diagram on right. The lower panel is a shorter exposure. G1-3, 4, 5 are ABCF3 knockout clonal lines and gSc-1, 2 are control scrambled knockout clonal lines.



**Fig S2 | GCN1 is required for GCN2 activation in HeLa.** shRNA-mediated scrambled and GCN1 knockdown cells were transfected with empty vector or shRNA-resistant GCN1 for 24 hours and then treated with indicated concentrations of HF for 1 hour. Lysates were assayed by immunoblot for phospho-GCN2 and phospho-eIF2α. Total GCN1 demonstrates efficient knockdown of GCN1 levels by shRNA and actin is shown as a loading control. Data are representative of a single experiment.



**Fig S3 | Distribution of ectopically expressed tagged GCN1 by immunoblot and fluorescence microscopy.** (a) Comparison of commercial GCN1 antibodies and comparison of tagged ectopically expressed GCN1 in digitonin sequential extractions by immunoblot in HeLa. c-Myc, erp29, eIF2α, and actin are provided as markers for cell fractions and as loading controls. (b) (top panel) Live-cell images of td-Tomato-GCN1 fluorescence in td-Tomato-GCN1 transfected HeLa. (bottom panel) Images of td-Tomato-GCN1 fluorescence in fixed td-Tomato-GCN1 transfected K4. Data are representative of a single experiment.

## REFERENCES

1. Grohmann U, *et al.* (2017) Amino-acid sensing and degrading pathways in immune regulation. *Cytokine & growth factor reviews* 35:37-45.
2. Cobbold SP, *et al.* (2009) Infectious tolerance via the consumption of essential amino acids and mTOR signaling. *Proceedings of the National Academy of Sciences of the United States of America* 106(29):12055-12060.
3. Zeng W, *et al.* (2011) A genetic screen reveals Arabidopsis stomatal and/or apoplastic defenses against *Pseudomonas syringae* pv. tomato DC3000. *PLoS pathogens* 7(10):e1002291.
4. Zhang Y, *et al.* (2008) GCN2-dependent phosphorylation of eukaryotic translation initiation factor-2alpha in Arabidopsis. *Journal of experimental botany* 59(11):3131-3141.
5. Magnusson LU, Farewell A, & Nystrom T (2005) ppGpp: a global regulator in *Escherichia coli*. *Trends in microbiology* 13(5):236-242.
6. Braeken K, Moris M, Daniels R, Vanderleyden J, & Michiels J (2006) New horizons for (p)ppGpp in bacterial and plant physiology. *Trends in microbiology* 14(1):45-54.
7. Haigis MC & Yankner BA (2010) The Aging Stress Response. *Molecular cell* 40:333-344.
8. Fontana L, Partridge L, & Longo VD (2010) Extending Healthy Life Span-From Yeast to Humans. *Science* 328:321-326.
9. Thomson AW, Turnquist HR, & Raimondi G (2009) Immunoregulatory functions of mTOR inhibition. *Nature reviews. Immunology* 9(5):324-337.
10. Murray PJ (2016) Amino acid auxotrophy as a system of immunological control nodes. *Nature immunology* 17(2):132-139.
11. Noss EH & Brenner MB (2008) The role and therapeutic implications of fibroblast-like synoviocytes in inflammation and cartilage erosion in rheumatoid arthritis. *Immunological Reviews* 223:252-270.
12. Grohmann U & Bronte V (2010) Control of immune response by amino acid metabolism. *Immunological Reviews* 236:243-264.
13. McGaha TL, *et al.* (2012) Amino acid catabolism: a pivotal regulator of innate and adaptive immunity. *Immunological Reviews* 249:135-157.
14. Keller TL, *et al.* (2012) Halofuginone and other febrifugine derivatives inhibit prolyl-tRNA synthetase. *Nature Chemical Biology* 8:311-317.
15. Zhou H, Sun L, Yang XL, & Schimmel P (2013) ATP-directed capture of bioactive herbal-based medicine on human tRNA synthetase. *Nature* 494(7435):121-124.
16. Sundrud MS, *et al.* (2009) Halofuginone inhibits Th17 cell differentiation by activating the amino acid starvation response. *Science* 324:1334-1338.
17. Stockinger B, Veldhoen M, & Martin B (2007) Th17 T cells: linking innate and adaptive immunity. *Seminars in immunology* 19(6):353-361.
18. Stockinger B & Veldhoen M (2007) Differentiation and function of Th17 T cells. *Current opinion in immunology* 19(3):281-286.
19. Filer A (2013) The fibroblast as a therapeutic target in rheumatoid arthritis. *Current opinion in pharmacology* 13(3):413-419.
20. Bottini N & Firestein GS (2013) Duality of fibroblast-like synoviocytes in RA: passive responders and imprinted aggressors. *Nature reviews. Rheumatology* 9(1):24-33.
21. Marton MJ, Crouch D, & Hinnebusch AG (1993) GCN1, a Translational Activator of *GCN4* in *Saccharomyces cerevisiae*, is required for Phosphorylation of Eukaryotic Translation Initiation Factor 2 by Protein Kinase GCN2. *Molecular and cellular biology* 13(6):3541-3556.
22. Vasquez de Aldana CR, Marton MJ, & Hinnebusch AG (1995) GCN20, a novel ATP binding cassette protein, and GCN1 reside in a complex that mediates activation of the eIF-2alpha kinase GCN2 in amino acid-starved cells. *The EMBO journal* 14(13):3184-3199.

23. Marton MJ, Vasquez de Aldana CR, Qiu H, Chakraborty K, & Hinnebusch AG (1997) Evidence that GCN1 and GCN20, Translational Regulators of *GCN4*, Function on Elongating Ribosomes in Activation of eIF2 $\alpha$  Kinase GCN2. *Molecular and cellular biology* 17(8):4474-4489.
24. Garcia-Barrio M, Dong J, Ufano S, & Hinnebusch AG (2000) Association of GCN1-GCN20 regulatory complex with the N-terminus of eIF2 $\alpha$  kinase GCN2 is required for GCN2 activation. *The EMBO journal* 19(8):1887-1899.
25. Sattlegger E & Hinnebusch AG (2005) Polyribosome binding by GCN1 is required for full activation of eukaryotic translation initiation factor 2 $\alpha$  kinase GCN2 during amino acid starvation. *The Journal of biological chemistry* 280(16):16514-16521.
26. Sattlegger E & Hinnebusch AG (2000) Separate domains in GCN1 for binding protein kinase GCN2 and ribosomes are required for GCN2 activation in amino acid-starved cells. *The EMBO journal* 19(23):6622-6633.
27. Kubota H, Ota K, Sakaki Y, & Ito T (2001) Budding yeast GCN1 binds the GI domain to activate the eIF2 $\alpha$  kinase GCN2. *The Journal of biological chemistry* 276(20):17591-17596.
28. Hirose T & Horvitz HR (2014) The translational regulators GCN-1 and ABCF-3 act together to promote apoptosis in *C. elegans*. *PLoS genetics* 10(8):e1004512.
29. Kaplan A, et al. (2017) Small-Molecule Stabilization of 14-3-3 Protein-Protein Interactions Stimulates Axon Regeneration. *Neuron* 93(5):1082-1093 e1085.
30. Pyper SR, et al. (2010) PRIC295, a Nuclear Receptor Coactivator, Identified from PPAR $\alpha$ -Interacting Cofactor Complex. *PPAR research* 2010.
31. Ebmeier CC & Taatjes DJ (2010) Activator-Mediator binding regulates Mediator-cofactor interactions. *Proceedings of the National Academy of Sciences* 107(25):11283-11288.
32. Shao W & Espenshade PJ (2012) Expanding roles for SREBP in metabolism. *Cell metabolism* 16(4):414-419.
33. Abramson J, Giraud M, Benoist C, & Mathis D (2010) Aire's partners in the molecular control of immunological tolerance. *Cell* 140(1):123-135.
34. Ewing RM, et al. (2007) Large-scale mapping of human protein-protein interactions by mass spectrometry. *Molecular systems biology* 3:89.
35. Koch HB, et al. (2007) Large-scale identification of c-MYC-associated proteins using a combined TAP/MudPIT approach. *Cell Cycle* 6(2):205-217.
36. Bandyopadhyay S, et al. (2010) A Human MAP Kinase Interactome. *Nature Methods* 7(10):801-805.
37. Jeronimo C, et al. (2007) Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. *Molecular cell* 27(2):262-274.
38. Lleres D, Denegri M, Biggiogera M, Ajuh P, & Lamond AI (2010) Direct interaction between hnRNP-M and CDC5L/PLRG1 proteins affects alternative splice site choice. *EMBO reports* 11(6):445-451.
39. Ajuh P, et al. (2000) Functional analysis of the human CDC5L complex and identification of its components by mass spectrometry. *The EMBO journal* 19(23):6569-6581.
40. Han SJ, et al. (2016) The Dual Estrogen Receptor  $\alpha$  Inhibitory Effects of the Tissue-Selective Estrogen Complex for Endometrial and Breast Safety. *Molecular pharmacology* 89(1):14-26.
41. Cambiaghi TD, et al. (2014) Evolutionarily conserved IMPACT impairs various stress responses that require GCN1 for activating the eIF2 kinase GCN2. *Biochemical and biophysical research communications* 443(2):592-597.
42. Herman JD, et al. (2015) The cytoplasmic prolyl-tRNA synthetase of the malaria parasite is a dual-stage target of febrilegugine and its analogs. *Science Translational Medicine* 7(288).
43. Gantke T, Sriskantharajah S, Sadowski M, & Ley SC (2012) I $\kappa$ B kinase regulation of the TPL-2/ERK MAPK pathway. *Immunological Reviews* 246:168-182.

44. Furusawa T, Moribe H, Kondoh H, & Higashi Y (1999) Identification of CtBP1 and CtBP2 as Corepressors of Zinc Finger-Homeodomain Factor  $\delta$ EF1. *Molecular and cellular biology* 19(12):8581-8590.
45. Dean M & Annilo T (2005) Evolution of the ATP-binding cassette (ABC) transporter superfamily in vertebrates. *Annual review of genomics and human genetics* 6:123-142.
46. Rossi A, *et al.* (2015) Genetic compensation induced by deleterious mutations but not gene knockdowns. *Nature* 524(7564):230-233.
47. El-Brolosy MA & Stainier DYR (2017) Genetic compensation: A phenomenon in search of mechanisms. *PLoS genetics* 13(7):e1006780.
48. Nakamura A & Kimura H (2017) A new role of GCN2 in the nucleolus. *Biochemical and biophysical research communications* 485(2):484-491.
49. David A, *et al.* (2012) Nuclear translation visualized by ribosome-bound nascent chain puromycylation. *The Journal of cell biology* 197(1):45-57.
50. Kramer EB & Hopper AK (2013) Retrograde transfer RNA nuclear import provides a new level of tRNA quality control in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences* 110(52):21042-21047.
51. Plevin MJ, Mills MM, & Ikura M (2005) The LxxLL motif: a multifunctional binding sequence in transcriptional regulation. *Trends in biochemical sciences* 30(2):66-69.
52. Folco EG, Lei H, Hsu JL, & Reed R (2012) Small-scale nuclear extracts for functional assays of gene-expression machineries. *Journal of visualized experiments : JoVE* (64).





## **CHAPTER FOUR**

### **Characterization of a Novel Mammalian Nutritional Stress Response Pathway**



## INTRODUCTION

In the previous chapter I determined that the evolutionarily conserved protein, GCN1, is required for GCN2-dependent AAR pathway signaling in mammalian cells. I also found that GCN1 is required for GCN2-independent changes in key regulators of arthritis tissue remodeling induced by HF treatment in TNF $\alpha$ -stimulated cells. To investigate more generally how GCN1 might mediate HF effects on transcriptional programs, I next used unbiased transcriptomic analysis to compare HF effects in wild type FLS versus GCN1-knockdown or GCN2-knockdown FLS. Although we know GCN1 mediates these GCN2-independent effects, we do not know whether GCN1 mediates additional GCN2-independent effects or functions, or whether GCN1 mediates GCN2-independent effects in the absence of an inflammatory signal. The previous chapter looked at GCN1 activity in the absence of an inflammatory signal in non-FLS cell types. In this chapter, we return to our *in vitro* FLS system, in which we know GCN1 regulates transcription of several genes independently of GCN2. With this system, we use transcriptomic analysis to characterize the entire transcriptional program mediated by GCN1 upon HF treatment in both the presence and absence of TNF $\alpha$  revealing a substantial portion of AAR transcriptional changes are regulated by GCN1 independently of GCN2 in the presence and absence of an inflammatory stimulus.

How GCN1 might mediate GCN2 independent effects remains unknown. In the previous chapter, I tested the possibility, suggested by prior literature, that GCN1 might interact with nuclear regulators, and whether HF could change the subcellular distribution of GCN1. My studies of GCN1 nuclear localization and interaction were inconclusive, but differential extraction studies suggested that HF might indeed change GCN1 distribution. I therefore sought: 1) to further examine HF effects on GCN1 distribution, focusing on the ribosomal fraction, in which GCN1 is known to function in yeast; 2) to use unbiased mass spec based interactomics to identify GCN1 interactors that can subsequently be tested as candidate mediators of HF signaling to the GCN1-regulated transcriptional program identified in my transcriptomic analysis. In my nuclear immunoprecipitation of endogenous GCN1 from HeLa cells (chapter 3), I found that commercially available GCN1 antibodies interacted non-specifically with nuclear proteins, and therefore significantly compromised experimental interpretation. Moving forward, I developed cells expressing a knocked-in GCN1 gene with an epitope tag to allow immuno-isolation studies with an anti-epitope antibody. This technical improvement allows future interactomic studies to be performed, while minimizing the issues of commercial antibodies to endogenous GCN1 cross-reacting non-specifically with cellular proteins. In these experiments, I will examine HF-treatment driven changes in GCN1 protein associations in TNF $\alpha$ -treated FLS, to better understand which of these changes can be linked causally to therapeutic inflammatory suppressive effects of the newly discovered nutritional stress response pathway.

## RESULTS

### **GCN1 is required for HF transcriptional regulation with a large portion independent of GCN2**

We know GCN1 mediates GCN2-dependent AAR pathway signaling and GCN2-independent inhibition of several therapeutic pro-inflammatory genes but we do not know what additional effects this branch point has or if the new branch functions outside an inflammatory context. To characterize the full transcriptional program regulated by GCN1 independently of GCN2 we compared changes in global mRNA expression between shRNA-mediated knockdown control, GCN1, and GCN2 K4 cells treated with and without HFol in the presence or absence of the inflammatory cytokine, TNF $\alpha$ . Fewer genes were differentially expressed upon HFol treatment in GCN1- and GCN2-knockdown cells compared to control knockdown cells consistent with a role for GCN1 and GCN2 in HF transcriptional regulation (**Table 1**). 7,521 genes were differentially expressed by HFol treatment in control cells compared to only 4,373 and 6,256 in GCN1- and GCN2-knockdown cells, respectively (**Table 1, Fig S1**). Similar levels of genes were significantly differentially expressed by TNF $\alpha$  treatment across cell types indicating GCN1 and GCN2 knockdown did not globally alter TNF $\alpha$  response (**Table 1, Fig S1 Fig 1c**).

Both in the presence and absence of TNF $\alpha$ , GCN1 was required for essentially all transcriptional regulation by HFol treatment including GCN2-dependent and -independent changes, while GCN2 was required for a smaller portion of these changes (**Fig 1a-b**). In GCN1-knockdown cells, there was little transcriptional regulation by HFol. A much smaller number of genes were differentially expressed upon HFol treatment in GCN1 knockdown cells compared to control cells and those genes that were differentially expressed exhibited a large decrease in magnitude in their fold change. The number of genes differentially expressed by HFol treatment decreased by almost 2 fold compared to control cells and the proportion of differentially expressed genes changing over 2 fold by HFol treatment fell from 18.7% to 1.2% (**Fig 1a-b, Table 1, Fig S1**). Similar changes occurred in the presence or absence of TNF $\alpha$  (**Fig 1a-b, Table 1, Fig S1**). In contrast, the number of differentially expressed genes upon HFol treatment in GCN2-knockdown cells compared to control cells was decreased by less than 1/5 compared to control cells in the absence of TNF $\alpha$  (**Fig 1a-b, Table 1, Fig S1**). Unexpectedly, in the presence of TNF $\alpha$ , the number of differentially expressed genes in GCN2-knockdown cells was greater than in control cells; however, this is consistent with previous observations by qPCR (**Table 1**, qPCR data not shown). These results confirm that GCN1 mediates both GCN2 dependent and independent transcriptional changes by HF, validating GCN1 as a branch point in the AAR. Notably, GCN1 mediates a large transcriptional program independent of GCN2 in both the presence and absence of TNF $\alpha$ .

In the presence of TNF $\alpha$ , previously identified GCN1-dependent, GCN2-independent genes such as CXCL10, MMP13, and IDO1 were regulated as expected i.e. only inhibited by HFol in control and GCN2-knockdown cells. Additional genes revealed within the transcriptomics as GCN1-dependent and GCN2-independent in the presence of TNF $\alpha$  included chemokines, cytokines, cell surface proteins (e.g. cadherin 5, type 2, lectin), transcriptional regulators (e.g. ZMYND15, BATF2), and genes involved in blood vessel development and ECM remodeling consistent with known therapeutic activities of HF. Notably, in the absence of TNF $\alpha$  stimulation, GCN1-dependent but GCN2-independent differentially expressed genes by HFol were significantly enriched in gene ontology terms related to cholesterol and steroid metabolic processes, while GCN2-dependent differentially expressed genes were not (**Table S2-3**).

### **HF treatment shifts GCN1 out of the ribosome fraction**

GCN1 clearly mediates a large set of GCN2-independent transcriptional effects but we still don't know how. In the previous chapter immunofluorescence analysis suggested that most if not all mammalian GCN1 resides in the cytoplasm. In yeast, GCN1 interacts with translating ribosomes and ribosomal association is required for GCN1 mediation of the uncharged tRNA signal to GCN2 (1, 2). To determine whether HF alters GCN1 association with ribosomes in mammals, we lysed untreated or HF treated K4 cells with digitonin to obtain a cytoplasmic cell fraction, and spun this fraction through a sucrose cushion to obtain a crude ribosomal pellet, which we then analyzed by immunoblot. As expected, ribosomal proteins L13a and L7a were concentrated in the pellet, while cytoplasmic proteins such as GAPDH were almost undetectable in the pellet (**Fig 2a**). HF treatment decreased GCN1 in the ribosomal pellet but did not alter ribosomal protein levels demonstrating that HF shifts GCN1 out of the ribosomal fraction. In yeast, ATP enhances GCN1 association with polysomes. Likewise, the presence of ATP increased GCN1 in the ribosomal pellet (**Fig 2a**). Notably GCN2 and ABCF3, whose orthologs are known to interact with GCN1 in yeast, exhibited patterns similar to GCN1 across ribosomal pellets, while HF's target, EPRS was not similarly distributed (**Fig 2a**). Cycloheximide treatment, which results in stalled ribosomes, increased GCN1 in the ribosomal pellet indicating that GCN1 in the ribosomal pellet is associated with the ribosome (**Fig S2**). In yeast, GCN1 physically interacts with the small ribosomal protein S10 (Rps10) and Rps10 mutant strains show reduced eIF2 $\alpha$  phosphorylation under nutrient replete conditions and within a few minutes of amino acid starvation suggesting this interaction is involved in GCN1 function (3). Mammalian GCN1 likewise co-precipitated with mammalian Rps10 and HF treatment decreased GCN1 interaction with Rps10 consistent with the shift of GCN1 out of the ribosomal fraction with HF

treatment (**Fig 2b**). Together this suggests that upon uncharged tRNA stimulation, GCN1 is released from the ribosome to mediate functions away from the ribosome.

### **Does HF change GCN1 interactions in cytoplasmic and/or ER/Nuclear fractions?**

We know GCN1 is released from the ribosome upon HF treatment, suggesting GCN1 distribution within the cell plays a role in mediating GCN1 branch functions. Since GCN1 was not observed in the nucleus, but putatively interacts with a set of proteins distinct from its GCN2 function and enriched in nuclear regulators, GCN1 likely mediates its functions by interacting with other proteins. To determine what GCN1 is interacting with, where these interactions are taking place, and whether these interactions are regulated by the AAR we set up an unbiased approach to pull down GCN1 and identify GCN1 protein interactors by TMT-MS analysis from different cell fractions under conditions that recapitulate the conditions in which we have characterized a full transcriptional readout of this novel branch.

### **Knock-in cells allow specific and efficient pull down of GCN1**

In the previous chapter, immunoprecipitation of endogenous GCN1 from nuclear extracts revealed that commercial GCN1 antibody was not specific enough to reliably pull down GCN1 to identify interactors. Instead, we generated HA-tagged GCN1 expressing K4 cells via CRISPR knock-in, to allow us to specifically immunoprecipitate GCN1 with HA antibody from the same cells and conditions in which we have now characterized a complete transcriptional program mediated by the GCN1 branch. The distribution of HA-GCN1 expression across sequentially extracted cell fractions and ribosomal pellets was similar to the distribution of endogenous GCN1, although changes with HF were less noticeable, suggesting tagged-GCN1 behaves similarly to endogenous GCN1 (**Fig S3**).

### **Identification of protein components of the GCN1-dependent, GCN2-independent pathway**

To determine protein interactors we treated knock-in K4 cells with and without HF in the presence of TNF $\alpha$  to recapitulate conditions under which we know the GCN1 mediates GCN2-independent functions. HA-GCN1 was immunoprecipitated from digitonin and NP/40 high salt fractions and sent for TMT-MS analysis, to identify and quantify all proteins interacting with GCN1 under these conditions. Silverstain of these bands confirmed several protein bands specific to GCN1 immunoprecipitation (**Fig 3**). A subset of these bands changed with HF treatment, suggesting these bands are GCN1 protein interactors regulated by HF treatment.

## **DISCUSSION**

### **GCN1 is a major mediator of uncharged tRNA effects**

In this chapter we use unbiased approaches to identify both downstream transcriptional effects and upstream protein mediators of the GCN2-independent GCN1 branch. In the previous chapter, we found that GCN1 represents a branch point in the mammalian AAR, required for GCN2 pathway activation and for several previously identified GCN2-independent transcriptional effects. Transcriptomic comparison of HF treated control, GCN1-, and GCN2-knockdown K4 confirm GCN1 is a branch point in the AAR, whose GCN2-independent branch is a major mediator of transcriptional changes triggered by uncharged tRNA. Our transcriptomics show that in both the presence and absence of TNF $\alpha$ , GCN1 is required for pretty much all HF transcriptional changes including GCN2-independent changes, and GCN2-dependent changes are only a fraction of the total changes mediated by GCN1. A subset of the transcriptional changes observed will need to be confirmed in independent experiments by qPCR. If GCN1 and GCN2 knockout K4 can be generated, dependence on GCN1 and GCN2 can be clearly evaluated. In addition, the broader set of GCN1-dependent, GCN2-independent HF-regulated genes identified here, including ones occurring in the absence of TNF $\alpha$  may allow the use of other cells to confirm and functionally study the pathway, such as HCT116, in which GCN1 and GCN2 knockouts have been generated. In addition, these genes will provide readouts for study of this new pathway outside the inflammatory context. This data reveals not only what functions this GCN2-independent branch mediates, but patterns and

characteristics of the regulated genes will provide clues for upstream regulators that function between GCN1 and these transcriptional changes. Comparison to transcriptional programs regulated by HF in other cell types will provide further clues into possible master regulators involved in the GCN2-independent branch across cell types.

### **How does GCN1 function?**

GCN1 clearly mediates important transcriptional effects independently of GCN2, but we do not know how. To begin to understand what is happening between GCN1 and its transcriptional effects we looked upstream of transcription, to whether HF treatment induced changes to GCN1 itself, in particular whether HF changed its localization and its protein interactors. We found that HF treatment both shifts GCN1 out of the ribosomal fraction and decreases its interaction with the ribosomal protein, Rps10. This suggests HF releases GCN1 from the ribosome, potentially allowing GCN1 to mediate functions away from the ribosome. We previously observed in sequential cell fractionations that HF treatment shifts GCN1 into the digitonin fraction and out from the NP-40 and high salt fractions. This would be consistent with the release of GCN1 from the ribosome upon increased levels of uncharged tRNA, which would include release of GCN1 into the digitonin-soluble cytoplasmic fraction from NP-40 soluble ER-bound ribosomes. Interestingly, ATP also increased GCN1 in the ribosomal pellet. In yeast, ATP increases GCN1 association with the ribosome in a GCN20-dependent fashion, although the functional significance is unknown (1). However, in the previous chapter we showed that ABCF3 was not required for HF-induced GCN2 pathway activation or transcriptional inhibition of several GCN2-independent genes by HF treatment. In addition, both ABCF3 and GCN2 were regulated by HF and ATP similarly to GCN1, suggesting they may form a complex with GCN1, as they do in yeast. What role ABCF3 plays, especially with respect to the findings of the previous chapter, will need to be further tested.

GCN1 likely interacts with currently unknown proteins to mediate its GCN2-independent functions. The optimized system introduced in this chapter will not only identify interactions in an unbiased manner, but will indicate whether HF regulates these interactions. If an interaction is changed by HF and independently confirmed, knockdown or knockout cell lines for the protein of interest can be used to determine whether the interaction is required for regulating transcriptional changes characterized earlier in this chapter. The known functions of identified proteins will also likely provide functional insight. Once these interactions are identified we can determine if there is any overlap with our putative list of nuclear GCN1 interactors from the last chapter or putative GCN1 interactors identified in previous studies (e.g. ABCF3, mediator complex). Our optimized system will further determine whether interactions occur in the digitonin-soluble cytoplasm associated fraction or the NP-40/high salt ER/nuclear associated fraction, which will indicate whether GCN1 localization regulates or promotes certain protein interactions as well as provide potential protein regulators of GCN1 localization. In addition, since GCN1 mediates GCN2-independent effects triggered by uncharged tRNA another uncharged tRNA sensor must exist. Either GCN1 is an uncharged tRNA sensor directly binding uncharged tRNA itself or GCN1 may function as a scaffold for a currently unidentified uncharged tRNA sensor. Protein interactors identified here will provide candidates for this function, which can be further tested. The non-canonical GCN1 branch plays a large role in the mammalian cellular response to uncharged tRNA, previously only thought to consist of the GCN2 pathway. This chapter identifies both downstream transcriptional effects and sets up a system to identify upstream mediators of this previously unidentified pathway to begin to understand its functions and mechanisms.

## METHODS

### Reagents

As described in chapter three. Additional reagents used in this chapter are listed below.

ANTIBODIES: L7a (CST 2415), L13a (CST 2765S), EPRS (Abcam ab31531)

### Cell Lines and Cell Culture

As described in chapter three.

### cDNAs

pCSX-HA-GCN1 (N-terminal tag):

pAAVS1-puro-DNR-HA-GCN1 (N-terminal tag): PmeI and NheI used to clone HA-GCN1 from pCSX-HA-GCN1 and into pAAVS1-puro-DNR (with previously modified multiple cloning site).

### Generation of HA-GCN1 K4 stable cell line

HA-GCN1 was knocked into the AAVS1 genomic locus in K4 cells by CRISPR using the AAVS1 Transgene knockin vector kit according to manufacturer's protocol (Origene GE100027). K4 cells were transiently transfected with pCas-Guide-AAVS1 and pAAVS1-puro-DNR-HA-GCN1 and then maintained through 8 passages to dilute out non-integrated puromycin-resistant plasmid DNA. Cells were then puromycin selected and propagated.

### RNAseq sample collection and preparation

shRNA-mediated scrambled, GCN1, and GCN2 knockdown K4 cells (described in chapter three) were generated and obtained from Yeon-Jin Kim. Cells were starved in 0.2% FBS K4 media for 24 hours. Cells were then treated with 250 nM HFol or equivalent volume of DMF for 24 hours with or without addition of 10 ng/mL TNF $\alpha$  8 hours prior to RNA harvest. Biological duplicates were collected for each condition for a total of 24 samples. RNA quantity and quality was checked using a Bioanalyzer at the Biopolymers Facility, Harvard Medical School. Poly-A enriched libraries were prepared with the TruSeq Stranded mRNA Library Prep Kit, sequenced using a single-end approach with 100 bp reads at a total sequencing depth of 30 million reads on an Illumina 4000 instrument, and mapped to the genome by the JP Sulzberger Columbia Genome Center. Pairwise DEseq2 analysis of differentially expressed genes was performed on the Galaxy web platform (4). Gene ontology term enrichment was performed using STRING, v10.5 (5)

### Sucrose density centrifugation to isolate ribosomal pellets

K4 cells grown on 10 cm dishes were sequentially extracted by digitonin and NP-40 as described in chapter three but with an additional 2.5 mM MgCl<sub>2</sub> added to the digitonin buffer, 20 U/mL of RNase inhibitor added to all lysis buffers, and lysed in a volume of 3 mL. Protease and phosphatase inhibitors were added to all buffers. 3 mL lysate was layered over a 1.5 mL 1M sucrose cushion (1 M sucrose, 110 mM KOAc, 25 mM KHEPES, pH 7.2, 5 mM MgCl<sub>2</sub>, 1 mM EGTA), and centrifuged at 250,000 G for 2 hours at 4C in a Beckman-Coulter Optima XE-100 Ultracentrifuge with the Beckman-Coulter SW 55 Ti Rotor. The pellet was washed with 4.5 mL wash buffer (110 mM KOAc, 25 mM KHEPES, pH 7.2, 5 mM MgCl<sub>2</sub>, 1 mM EGTA, 100 mM sucrose) and centrifuged at 250,000G for another 30 minutes. The supernatant was carefully removed and the ribosomal pellet was re-suspended in 100  $\mu$ L 0.1% NP-40 buffer (0.1% NP-40, 110 mM KOAc, 25 mM KHEPES, pH 7.2, 5 mM MgCl<sub>2</sub>, 1 mM EGTA).

### Rps10 Co-IP

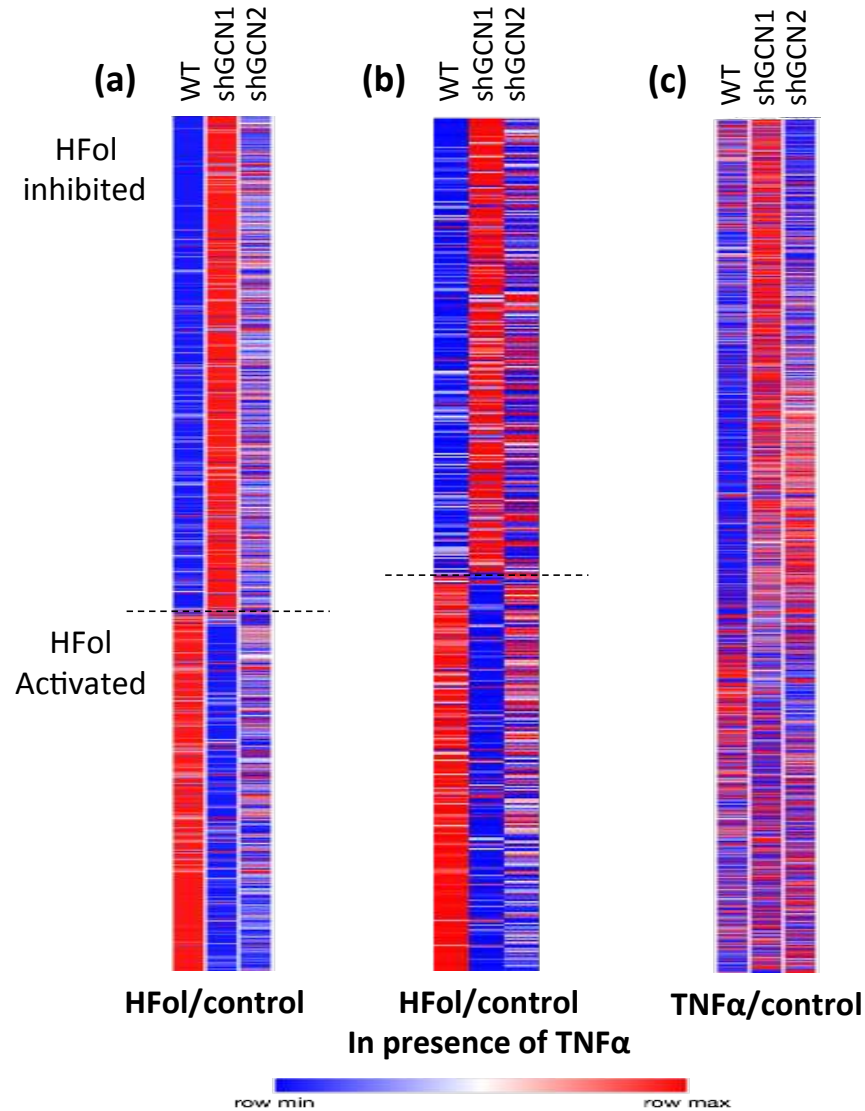
HA-GCN1 knockin K4 cells were treated with or without 500 nM HF for one hour and then all cells were treated with 100  $\mu$ g/mL cycloheximide for 4 minutes. Cells were digitonin permeabilized as described above with RNase inhibitor and ATP. After lysis NP-40 was added to lysate for a final concentration of 0.1%. HA immunoprecipitation and elution were performed as described below.

### **GCN1 pull down for TMT-MS**

HA-GCN1 knockin K4 cells grown on 10 cm dishes were starved in 0.2% FBS K4 media for 24 hours. All cells were then treated with 10 ng/mL TNF $\alpha$ . After 3 hours, 500 nM HF was added to HF treated cells, and after another 26 minutes all cells were treated with 100  $\mu$ g/mL cycloheximide for 4 minutes and then sequentially extracted by digitonin and NP-40/high salt as previously described with the following modification. For NP-40/high salt lysis, cells were scraped in 1 mL NP-40/high salt lysis buffer into microcentrifuge tubes, incubated for 10 minutes at 4C, centrifuged to remove cellular debris, and diluted with 4 mL dilution buffer (1% NP-40, 50 mM KHEPES, pH 7.2, 30 mM MgCl<sub>2</sub>). Lysates were incubated with magnetic HA beads (Pierce) for 3 hours at 4C in siliconized tubes. Beads pre-incubated with HA peptide were used as control. Beads were washed 3x with lysis buffer and 1x with 0.1% NP-40 PBS. Four 10 cm plates were pooled for each condition and each condition was performed in duplicate. Beads were eluted with PGB without bromophenol blue for 20 minutes at room temperature as described for the nuclear IP-MS in chapter three. 10% of each sample was run on the gel in figure 3 for silverstaining and the rest was sent to the Thermo Fisher Center for Multiplexed Proteomics, Harvard Medical School for quantitative protein identification by TMT-MS analysis.



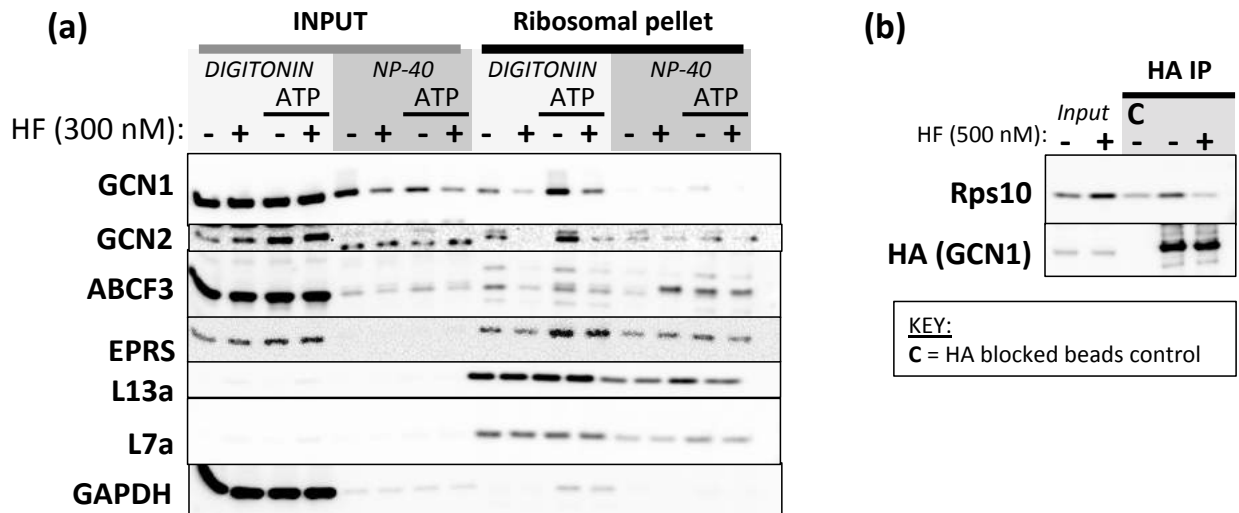
## FIGURES AND TABLES



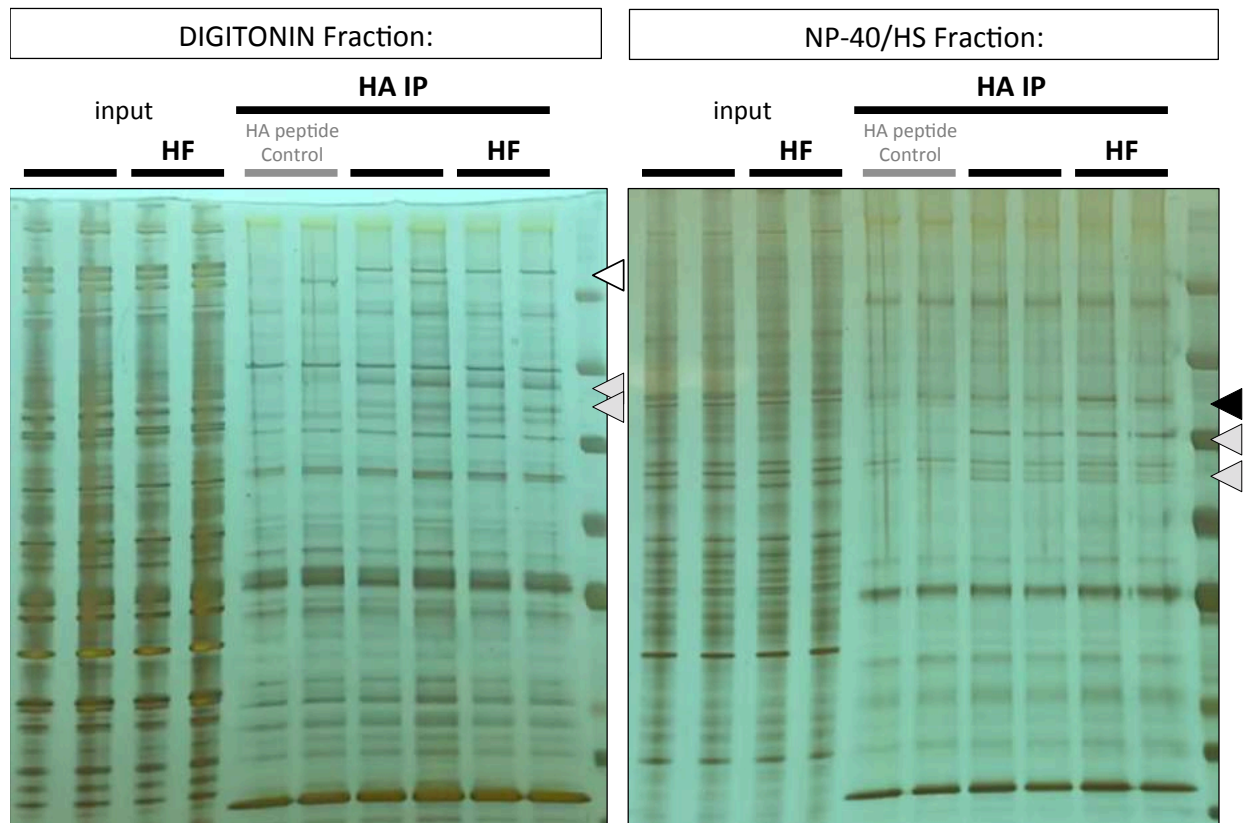
**Fig 1 | GCN1 is required for most if not all HFol regulated mRNA expression, while GCN2 is required for only a portion of HFol regulated mRNA expression.** (a) Heatmap of average fold change by HFol treatment relative to untreated in shRNA mediated control, GCN1, and GCN2 knockdown K4. (b) Heatmap of average fold change by HFol treatment with TNF $\alpha$  stimulation relative to TNF $\alpha$  treatment alone. (c) Heatmap of average fold change by TNF $\alpha$  treatment relative to untreated. In all heatmaps only genes with values that changed over 2.5 fold in control cells were mapped.

Cell type	Differentially expressed genes	Differentially expressed genes with change > 2 fold
<b>HFol treated versus untreated</b>		
control	7521	1409
shGCN1	4373	52
shGCN2	6256	808
<b>TNF<math>\alpha</math> treated versus untreated</b>		
control	6899	1792
shGCN1	6354	1374
shGCN2	6190	1625
<b>HFol/TNF<math>\alpha</math> treated versus TNF<math>\alpha</math> treated</b>		
control	7603	1362
shGCN1	4355	54
shGCN2	8140	1904

**Table 1 | Significantly differentially expressed genes.** Genes were considered differentially expressed if  $p_{Adj} < 0.05$ . Control is shRNA-mediated scrambled knockdown K4, shGCN1 is shRNA-mediated GCN1 knockdown K4, and shGCN2 is shRNA-mediated GCN2 knockdown K4.

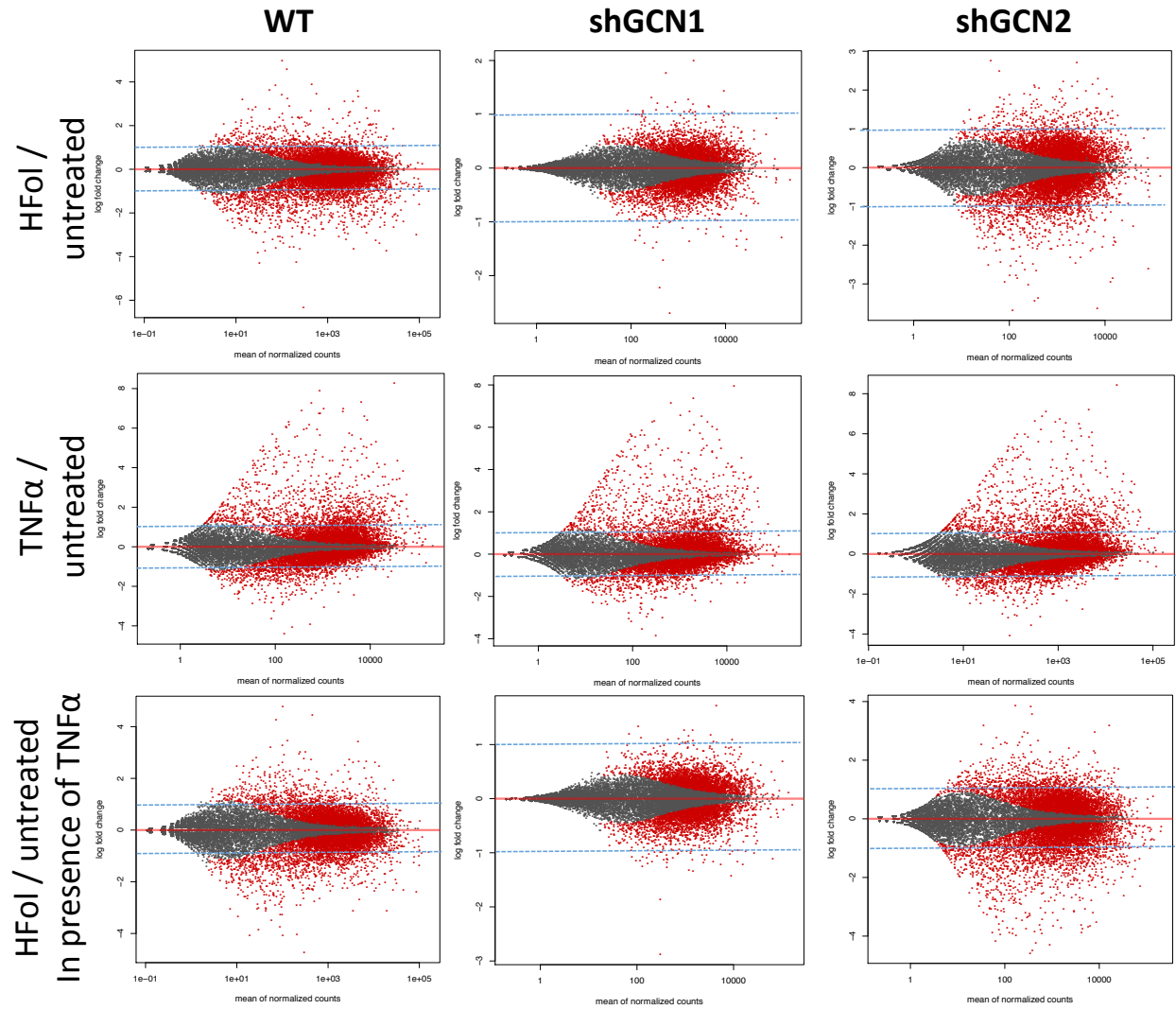


**Fig 2 | HF treatment shifts GCN1 out of ribosome fraction and decreases GCN1 interaction with the small ribosomal protein, Rps10. (a)** Digitonin and NP-40 extracts from K4 cells treated with and without HF and harvested in the presence or absence of 5 mM ATP were layered over a sucrose cushion and centrifuged at high speed. The pellet was re-suspended and analyzed by immunoblot for the indicated proteins. **(b)** GCN1 from stable HA-GCN1 knockin K4 was immunoprecipitated by HA antibody-linked beads and assayed for Rps10 and HA-GCN1 protein by immunoblot. HA peptide blocked beads were used as control. HF treatment data in (a) is representative of at least 3 separate experiments. ATP treatment compared to no ATP treatment was tested only in this single experiment. Data in (b) is representative of 3 separate experiments.



**Fig 3 | Optimized unbiased IP-MS approach to identify GCN1 interactors shows proteins specific to GCN1 pulldown changed by HF treatment.** Silverstained protein gels comparing proteins precipitated by epitope tag pull down of GCN1 from non-treated and HF treated HA-GCN1 knockin K4 in the presence of  $\text{TNF}\alpha$ . Cells were sequentially fractionated, first by digitonin, and then by NP-40/high salt, and GCN1 immunoprecipitated from each fraction (left panel) Digitonin fraction. (right panel) NP-40/high salt fraction. White arrow head indicates band at size of full length GCN1. Grey arrow heads indicate bands specific to GCN1 pull down, while black arrow head indicates a band appearing only with HF treatment.

## SUPPLEMENTARY FIGURES AND TABLES



**Fig S1 | MA plots for pairwise DESEQ2 analyses.** Fold changes of all genes plotted against normalized expression. Red dots show genes whose fold changes were statistically significant. Note y-axes display different ranges. Dashed blue lines mark 1 and -1 on the y-axis for orientation.

pathway description	observed gene	false discovery
	count	rate
cholesterol biosynthetic process	10	7.52E-08
cholesterol metabolic process	13	7.71E-07
immune system process	48	1.58E-05
defense response	38	4.60E-05
immune response	36	7.18E-05
response to stress	65	0.000114
immune effector process	19	0.00018
steroid metabolic process	14	0.000206
organic hydroxy compound metabolic process	20	0.000206
receptor-mediated endocytosis	12	0.000975
regulation of apoptotic process	34	0.000975
intrinsic apoptotic signaling pathway in response to ER stress	6	0.000975
regulation of cell death	35	0.00124
organic hydroxy compound biosynthetic process	11	0.00196
small molecule biosynthetic process	16	0.00206
innate immune response	26	0.00206
alcohol metabolic process	15	0.00252
response to stimulus	98	0.00361
defense response to virus	10	0.0038
type I interferon signaling pathway	7	0.0041
cellular response to type I interferon	7	0.0041
response to external stimulus	39	0.00421
vasculature development	17	0.00474
apoptotic process	27	0.0053
single-organism process	133	0.0053
single-organism cellular process	128	0.00602
regulation of response to stimulus	57	0.00604
positive regulation of biological process	76	0.00781
blood vessel development	16	0.00842
positive regulation of response to stimulus	38	0.0109
response to biotic stimulus	20	0.0137
response to unfolded protein	9	0.0141
programmed cell death	26	0.0141
negative regulation of cell proliferation	19	0.0149
response to other organism	19	0.0214
response to nitrogen compound	22	0.0222
isoprenoid biosynthetic process	4	0.0309
negative regulation of interleukin-2 secretion	2	0.0343
positive regulation of I-kappaB kinase/NF-kappaB signaling	9	0.0349
positive regulation of cysteine-type endopeptidase activity		
involved in apoptotic signaling pathway	4	0.0349
anatomical structure development	63	0.0356
activation of immune response	14	0.0386
cellular response to interferon-alpha	3	0.0386
response to organonitrogen compound	20	0.0403
cell communication	73	0.0405
endoplasmic reticulum unfolded protein response	7	0.0405
system development	56	0.0438
signal transduction	68	0.0462
cellular response to unfolded protein	7	0.0462
positive regulation of apoptotic process	16	0.0462
cellular response to biotic stimulus	8	0.0488
regulation of blood pressure	8	0.0498

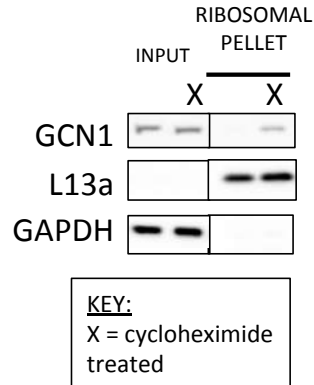
**Table S1 | Functional enrichments of most HF regulated genes in the absence of TNF $\alpha$ .** Genes with over 2 fold change in expression with HF treatment compared to non treated in control cells.

pathway description	observed gene false discovery	
	count	rate
cholesterol biosynthetic process	7	1.51E-06
cholesterol metabolic process	9	1.82E-06
steroid metabolic process	10	5.11E-05
vasculature development	10	0.0151
organic hydroxy compound metabolic process	10	0.0166
blood vessel development	9	0.0479
isopentenyl diphosphate biosynthetic process, mevalonate pathway	2	0.0479

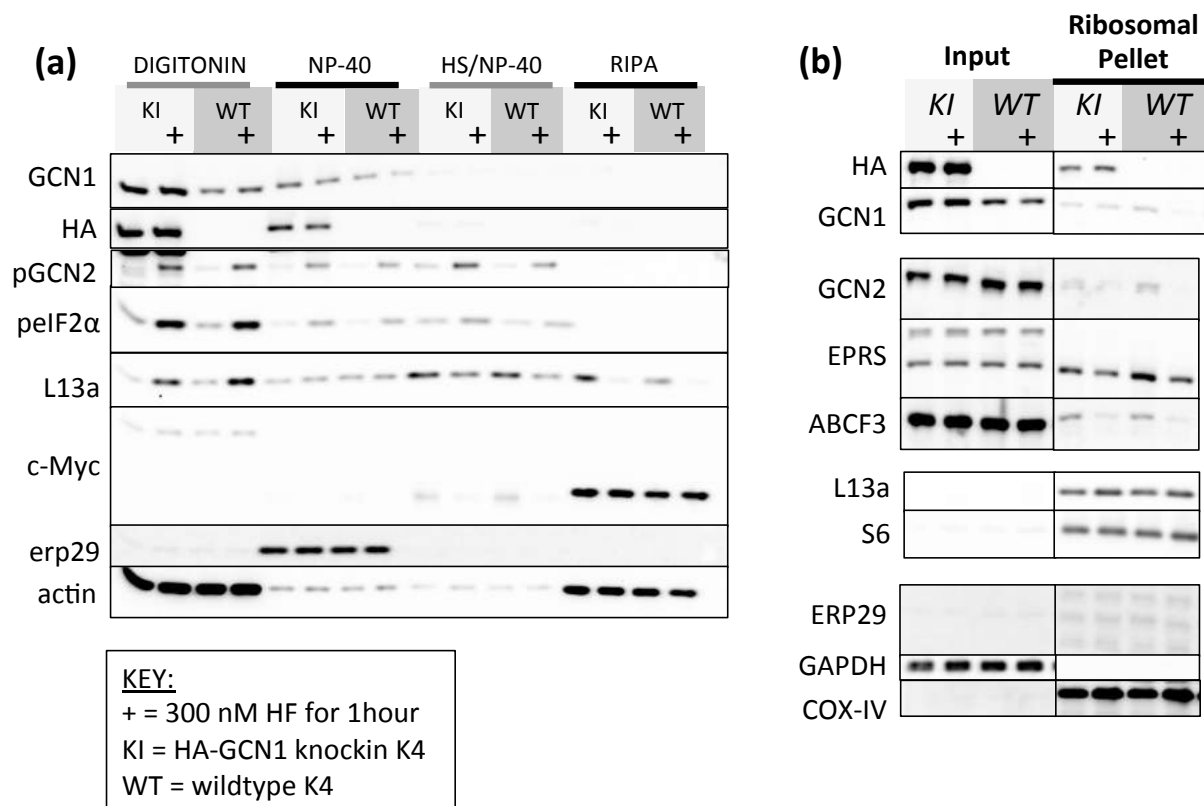
**Table S2 | Functional enrichments of GCN1 dependent, GCN2 independent genes in the absence of TNF $\alpha$ .** From genes with over 2 fold change in expression with HF treatment compared to non treated in control cell, designated them as GCN1 dependent/GCN2 independent if there was over a 2 fold change between the fold change in control versus shGCN1 knockdown cells, and over a 2 fold change in shGCN1 knockdown cells versus shGCN2 knockdown cells. About 88 genes.

pathway description	observed gene false discovery	
	count	rate
defense response	27	1.69E-05
immune system process	31	4.78E-05
immune response	24	0.00015
immune effector process	14	0.00017
type I interferon signaling pathway	7	0.00017
cellular response to type I interferon	7	0.00017
defense response to virus	9	0.000274
response to stimulus	59	0.00075
innate immune response	18	0.00163
response to stress	37	0.00207
response to other organism	15	0.00207
cytokine-mediated signaling pathway	11	0.00446
response to external stimulus	24	0.0135
cellular response to interferon-alpha	3	0.0147
interferon-gamma-mediated signaling pathway	5	0.0186
negative regulation of interleukin-2 secretion	2	0.0209
signal transduction	41	0.0363
cell surface receptor signaling pathway	24	0.0454
cellular response to stimulus	48	0.046
cell communication	43	0.0473
cellular response to molecule of bacterial origin	6	0.0481
cellular response to cytokine stimulus	11	0.0481

**Table S3 | Functional enrichments of GCN2 dependent genes in the absence of TNF $\alpha$ .** From genes with over 2 fold change in expression with HF treatment compared to non treated in control cell, designated them as GCN2 dependent if there was over a 2 fold change between the fold change in control versus shGCN2 knockdown cells. About 130 genes.



**Fig S2 | Cycloheximide treatment increases GCN1 levels in the ribosome fraction.** K4 cells treated with and without 100 ug/mL cycloheximide for 4 minutes were digitonin lysed in the presence of ATP. The lysate was layered over a sucrose cushion and centrifuged at high speed. The pellet was re-suspended and analyzed by immunoblot for the indicated proteins. Data are representative of a single experiment.



**Fig S3 | HA-GCN1 is distributed similarly to endogenous GCN1 by biochemical fractionation and sucrose cushion centrifugation.** (a) Tagged and endogenous levels of GCN1 were compared by immunoblot across digitonin sequential extraction between wildtype K4 and stable HA-GCN1 knockin K4 treated with and without HF. (b) Tagged and endogenous levels of GCN1 were compared between ribosomal pellets of wildtype K4 and stable HA-GCN1 knockin K4 treated with and without HF. Data are representative of a single experiment.

## REFERENCES

1. Marton MJ, Vasquez de Aldana CR, Qiu H, Chakraborty K, & Hinnebusch AG (1997) Evidence that GCN1 and GCN20, Translational Regulators of *GCN4*, Function on Elongating Ribosomes in Activation of eIF2 $\alpha$  Kinase GCN2. *Molecular and cellular biology* 17(8):4474-4489.
2. Sattlegger E & Hinnebusch AG (2005) Polyribosome binding by GCN1 is required for full activation of eukaryotic translation initiation factor 2 $\alpha$  kinase GCN2 during amino acid starvation. *The Journal of biological chemistry* 280(16):16514-16521.
3. Lee SJ, Swanson MJ, & Sattlegger E (2015) Gcn1 contacts the small ribosomal protein Rps10, which is required for full activation of the protein kinase Gcn2. *The Biochemical journal* 466(3):547-559.
4. Afgan E, *et al.* (2016) The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res* 44(W1):W3-W10.
5. Szklarczyk D, *et al.* (2017) The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res* 45(D1):D362-D368.



## **CHAPTER FIVE**

### **Conclusions**



## CONCLUSIONS

This thesis describes the discovery of a novel nutritional stress response pathway in mammalian cells that utilizes components of the AAR pathway upstream of GCN2, including the scaffold-like protein GCN1. The AAR pathway is part of the larger ISR in mammals, which consists of multiple branches triggered by distinct stresses, which share phospho-eIF2 $\alpha$  signal generation, but each also has a unique cellular response due to activation of parallel pathway(s) by the initiating stress (1-5). No parallel pathway has been identified for the AAR until now. Our observations are the first demonstration of an alternate, parallel pathway triggered by the accumulation of uncharged tRNA. The first part of my thesis reveals how conserved components of this stress response are, by demonstrating that core components of the mammalian ISR were already present over 550 million years ago in the common ancestor of cnidarians and bilaterians, playing a key role in stress adaptation not only in mammals but across metazoans. By looking at the conservation of the oldest part of the ISR, the AAR, evolutionary conservation helped us identify GCN1 as a mediator of an unidentified pathway induced by uncharged tRNA in mammalian cells observed by our lab. GCN1 appears conserved throughout eukaryotes, suggesting it is at least as old as GCN2, suggesting the GCN2-independent GCN1 stress response pathway described here may be conserved throughout eukaryotes. These observations as well as additional observations from the literature spanning eukaryotes, suggests the GCN2-independent GCN1 pathway mediates a conserved role in organismal defense.

### **GCN1 senses the uncharged tRNA signal without GCN2**

The capacity of the signal initiating apparatus of the AAR to function in the absence of GCN2, strongly suggests that GCN1 or a complex containing GCN1 can recognize, bind and signal in response to uncharged tRNA. In yeast, GCN1 is known to function at the ribosome and its ribosomal association is required for uncharged tRNA activation of GCN2 (6, 7). This led to the model in yeast that GCN1 detects the uncharged tRNA signal, either by direct binding of uncharged tRNA or by some other mechanism, at the ribosome (6, 8). Data shown here suggests that mammalian GCN1 likewise binds the ribosome and that its ribosomal association is regulated by uncharged tRNA, suggesting mammalian GCN1 as in the yeast model detects the uncharged tRNA signal at the ribosome. The most highly conserved part of GCN1 is the central portion of GCN1, which shows high similarity to eEF3, an ABC protein found only in yeast (8, 9). Fungal eEF3 facilitates release of uncharged tRNA from the ribosomal E-site upon delivery of charged tRNA at the A-site during translation (8). Although it is unknown if eEF3 directly binds uncharged tRNA, this similarity has been suggested to indicate that GCN1 plays a similar role with respect to uncharged tRNA at the ribosome. In yeast, mutation of the eEF3-like domain of GCN1 impairs GCN2 activation more than expected from its effects on GCN1 ribosome binding alone, suggesting this domain has essential effector function in GCN2 activation by uncharged tRNA, possibly as a domain that binds uncharged tRNA, but this remains to be tested (7). Rescue of knockdown or knockout GCN1 cells with various mutant GCN1 constructs could provide valuable insight into sequences important for GCN1 function. Direct binding experiments could also be attempted but additional components could be required for GCN1 uncharged tRNA binding, which could lead to false negatives in *in vitro* binding assays; however, a positive result would be informative. It is also possible that rather than acting as a direct binder of uncharged tRNA, GCN1 interacts with or scaffolds another currently unidentified uncharged tRNA sensor. The TMT-MS interactomics described in chapter four will likely yield insight into these questions by identifying GCN1 interactors, which may function directly as uncharged tRNA sensors or form a complex required for GCN1 sensing of uncharged tRNA.

### **Observations of GCN2-independence may represent observations of the GCN1 branch**

Whereas anecdotal literature has reported various cellular responses to amino acid limitation that occur without GCN2 from yeast to mammals, there has been no molecular explanation for these observations. The identified GCN1 branch could account for some of these observations. In yeast, GCN2 independent splicing of ribosomal transcripts is induced by amino acid starvation within minutes (10). Transcriptional

profiling in mammals comparing wildtype and GCN2 null cells upon amino acid deprivation reveals a significant set of transcriptional changes that are independent of GCN2 (11). Regulation of genes involved in amino acid biosynthetic pathways appeared largely dependent on GCN2, while amino acid regulation of genes involved in other processes such as cholesterol biosynthesis were independent of GCN2 (11). Interpreting what specific molecular mechanisms are responsible for these effects are difficult, since amino acid deprivation stimulates multiple pathways including mTORC1. Moreover, our data reveals that by treatment with an aaRS inhibitor, a large portion of transcriptional regulation by uncharged tRNA is in fact GCN2-independent. Interestingly, our transcriptomic analyses suggested GCN2-independent GCN1-dependent genes were enriched in cholesterol biosynthetic processes as well. The existence of a novel GCN1 pathway will need to be considered in studies that have solely considered GCN2-dependence as a proxy for uncharged tRNA signaling in the cell. In *A. thaliana*, a chemical known to prime the plant immune system and augment plant defense was found to bind to and inhibit activity of aspartyl-tRNA-synthetase (12). Intriguingly, enhanced immunity by this chemical was independent of GCN2 (12). It will be interesting to determine whether GCN2-independent effects in plants and yeast are likewise GCN1-dependent, which would indicate the GCN2-independent GCN1 branch may be conserved throughout eukaryotes.

### **The GCN1 branch may be a conserved protective uncharged tRNA signaling pathway**

Genetics in yeast, nematodes, and plants have suggested a role for GCN1 in a variety of regulatory events independent of GCN2, including responses to pathogens. Consideration of these genetic data in light of our findings points to a conserved role for a GCN1-dependent, GCN2-independent signaling downstream of amino acid limitation. Evidence for a functional role of GCN1 separate from GCN2 has been found in the nematode, *C. elegans*. In *C. elegans*, GCN1 and ABCF3, physically interact, and are both required for apoptosis during development and in response to ionizing radiation, while GCN2 is not (13). In yeast, GCN1 and GCN20 have genetic interactions distinct from GCN2 (14-16). Further support for additional GCN1 functions come from the range of putative yeast and mammalian GCN1 protein interactors identified, which include a particularly high number of nuclear regulators, seemingly distinct from GCN1's functional role with GCN2 (17-27). Notably many studies across taxa have implicated GCN1 in organismal defense, but whether GCN2 also plays a role hasn't been tested. In the plant, *A. thaliana*, GCN1 is required for non-host and basal resistance against the bacteria, *Pseudomonas syringae*, and a separate study found both GCN1 and ABCF3 are required for bacteria-triggered stomatal closure (28, 29). Interestingly, GCN1 plays a role in plant cold response, and in mammals and yeast, translational inhibition that allows resistance to cold stress has been found to be independent of GCN2, peIF2 $\alpha$ , and TOR (30). In yeast and *C. elegans* GCN1 and GCN20/ABCF3 play roles in DNA damage (13, 31). In mammals, the AAR has further evolved protective roles in the regulation of immunity (32-34). A central question that emerges with the identification of this new GCN1 pathway is how conserved is this novel branch, and if it is what are its roles.

## **FUTURE DIRECTIONS**

### **Identification of GCN1 binding proteins**

How this new GCN1-dependent, GCN2-independent pathway mediates its effects remains unknown. As described in this thesis we do know that GCN1 is involved upstream in the pathway and we have now identified a transcriptional program downstream of the pathway. Identification of GCN1 binding proteins will provide candidate mediators regulating the therapeutic transcriptional programs identified by transcriptomic analysis in chapter four. HF treatment of cells in which identified proteins have been knocked down or knocked out with subsequent analysis of transcriptional changes by qPCR will begin to reveal what role these interactions play. The identity of interacting proteins themselves may provide insight into further downstream mechanisms mediating transcriptional regulation. Patterns discovered within the gene set regulated by this pathway identified by the transcriptomics will provide candidate transcriptional mediators that might be regulated by the pathway, which can then be further tested.

## REFERENCES

1. Taniuchi S, Miyake M, Tsugawa K, Oyadomari M, & Oyadomari S (2016) Integrated stress response of vertebrates is regulated by four eIF2 $\alpha$  kinases. *Scientific reports* 6:32886.
2. Donnelly N, Gorman AM, Gupta S, & Samali A (2013) The eIF2 $\alpha$  kinases: their structures and functions. *Cellular and molecular life sciences : CMLS* 70(19):3493-3511.
3. Wek RC & Cavener DR (2007) Translational control and the unfolded protein response. *Antioxidants & redox signaling* 9(12):2357-2371.
4. Jacobs BL & Langland JO (1996) When Two Strands Are Better Than One: The Mediators and Modulators of the Cellular Responses to Double-Stranded RNA. *Virology* 219:339-349.
5. Furuyama K, Kaneko K, & Vargas V. PD (2007) Heme as a Magnificent Molecule with Multiple Missions: Heme Determines Its Own Fate and Governs Cellular Homeostasis. *Journal of Experimental Medicine* 213:1-16.
6. Marton MJ, Vasquez de Aldana CR, Qiu H, CHakraborty K, & Hinnebusch AG (1997) Evidence that GCN1 and GCN20, Translational Regulators of *GCN4*, Function on Elongating Ribosomes in Activation of eIF2 $\alpha$  Kinase GCN2. *Molecular and cellular biology* 17(8):4474-4489.
7. Sattlegger E & Hinnebusch AG (2005) Polyribosome binding by GCN1 is required for full activation of eukaryotic translation initiation factor 2 $\{\alpha\}$  kinase GCN2 during amino acid starvation. *The Journal of biological chemistry* 280(16):16514-16521.
8. Visweswarai J, Lee SJ, Hinnebusch AG, & Sattlegger E (2012) Overexpression of eukaryotic translation elongation factor 3 impairs Gcn2 protein activation. *The Journal of biological chemistry* 287(45):37757-37768.
9. Marton MJ, Crouch D, & Hinnebusch AG (1993) GCN1, a Translational Activator of *GCN4* in *Saccharomyces cerevisiae*, is required for Phosphorylation of Eukaryotic Translation Initiation Factor 2 by Protein Kinase GCN2. *Molecular and cellular biology* 13(6):3541-3556.
10. Pleiss JA, Whitworth GB, Bergkessel M, & Guthrie C (2007) Rapid, transcript-specific changes in splicing in response to environmental stress. *Molecular cell* 27(6):928-937.
11. Deval C, *et al.* (2009) Amino acid limitation regulates the expression of genes involved in several specific biological processes through GCN2-dependent and GCN2-independent pathways. *The FEBS journal* 276(3):707-718.
12. Luna E, *et al.* (2014) plant perception of b-aminobutyric acid is mediated by an aspartyl-trna synthetase. *Nature Chemical Biology* 10:450-456.
13. Hirose T & Horvitz HR (2014) The translational regulators GCN-1 and ABCF-3 act together to promote apoptosis in *C. elegans*. *PLoS genetics* 10(8):e1004512.
14. Brice C, Sanchez I, Bigey F, Legras J-L, & Blondin B (2014) A genetic approach of wine yeast fermentation capacity in nitrogen-starvation reveals the key role of nitrogen signaling. *BMC genomics* 15(495).
15. Costanzo M, *et al.* (2016) A global genetic interaction network maps a wiring diagram of cellular function. *Science* 353:aaf1420.
16. Costanzo M *et al.* (2010) The Genetic Landscape of a Cell. *Science* 327:425-431.
17. Shao W, Zumer K, Fujinaga K, & Peterlin BM (2016) FBXO3 Protein Promotes Ubiquitylation and Transcriptional Activity of AIRE (Autoimmune Regulator). *The Journal of biological chemistry* 291(34):17953-17963.
18. Abramson J, Giraud M, Benoist C, & Mathis D (2010) Aire's partners in the molecular control of immunological tolerance. *Cell* 140(1):123-135.
19. Ewing RM, *et al.* (2007) Large-scale mapping of human protein-protein interactions by mass spectrometry. *Molecular systems biology* 3:89.
20. Koch HB, *et al.* (2007) Large-scale identification of c-MYC-associated proteins using a combined TAP/MudPIT approach. *Cell Cycle* 6(2):205-217.
21. Bandyopadhyay S, *et al.* (2010) A Human MAP Kinase Interactome. *Nature Methods* 7(10):801-805.

22. Jeronimo C, *et al.* (2007) Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. *Molecular cell* 27(2):262-274.
23. Lleres D, Denegri M, Biggiogera M, Ajuh P, & Lamond AI (2010) Direct interaction between hnRNP-M and CDC5L/PLRG1 proteins affects alternative splice site choice. *EMBO reports* 11(6):445-451.
24. Ajuh P, *et al.* (2000) Functional analysis of the human CDC5L complex and identification of its components by mass spectrometry. *The EMBO journal* 19(23):6569-6581.
25. Ebmeier CC & Taatjes DJ (2010) Activator-Mediator binding regulates Mediator-cofactor interactions. *Proceedings of the National Academy of Sciences* 107(25):11283-11288.
26. Pyper SR, *et al.* (2010) PRIC295, a Nuclear Receptor Coactivator, Identified from PPARalpha-Interacting Cofactor Complex. *PPAR research* 2010.
27. Han SJ, *et al.* (2016) The Dual Estrogen Receptor alpha Inhibitory Effects of the Tissue-Selective Estrogen Complex for Endometrial and Breast Safety. *Molecular pharmacology* 89(1):14-26.
28. Monaghan J & Li X (2010) The HEAT repeat protein ILITYHIA is required for plant immunity. *Plant & cell physiology* 51(5):742-753.
29. Zeng W, *et al.* (2011) A genetic screen reveals Arabidopsis stomatal and/or apoplastic defenses against *Pseudomonas syringae* pv. tomato DC3000. *PLoS pathogens* 7(10):e1002291.
30. Wang L, *et al.* (2017) The inhibition of protein translation mediated by AtGCN1 is essential for cold tolerance in *Arabidopsis thaliana*. *Plant, cell & environment* 40(1):56-68.
31. Leung GP, Aristizabal MJ, Krogan NJ, & Kobor MS (2014) Conditional Genetic Interactions of RTT107, SLX4, and HRQ1 Reveal Dynamic Networks upon DNA Damage in *S. cerevisiae*. *G3: Genes | Genomes | Genetics* 4:1059-1069.
32. Cobbold SP, *et al.* (2009) Infectious tolerance via the consumption of essential amino acids and mTOR signaling. *Proceedings of the National Academy of Sciences of the United States of America* 106(29):12055-12060.
33. Grohmann U & Bronte V (2010) Control of immune response by amino acid metabolism. *Immunological Reviews* 236:243-264.
34. McGaha TL, *et al.* (2012) Amino acid catabolism: a pivotal regulator of innate and adaptive immunity. *Immunological Reviews* 249:135-157.